Roskilde University

The Role of CDX2 in Colon Cancer Development and Progression

Davidsen, Johanne

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Davidsen, J. (2020). *The Role of CDX2 in Colon Cancer Development and Progression*. Roskilde Universitet.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact rucforsk@kb.dk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

The Role of CDX2 in Colon Cancer Development and Progression

PhD Thesis – November 2020

Johanne Davidsen

Department of Science and Environment, Roskilde University

Center for surgical science, Zealand University Hospital

Supervisor: Professor Jesper T. Troelsen

Co-supervisor: Professor Ismail Gogenür

This thesis has been submitted to the Doctoral School of Science and Environment,

Roskilde University

Preface and acknowledgements

This thesis is submitted to the PhD school at the Department of Science and Environment at Roskilde University, Denmark in partial fulfillment of the requirements to obtain a PhD degree. The work presented here was carried out from July 2016 to June 2020 (with leave of absence due to maternity leave), in the laboratory of Professor Jesper T. Troelsen at Department of Science and Environment at Roskilde University, including one month, October 2019, at Kevin Myants laboratory at Institute of Genetics and Molecular Medicine, University of Edinburgh, Scotland.

The PhD project was supported by grants from the Independent Research Fund Denmark and the Region Zealand. Travel expenses were funded by the Clausen and Hustrus Fond.

First and foremost, I would like to thank my supervisor Jesper T. Troelsen for giving me the opportunity to work in the Gastro group at Roskilde University for the last six years. Thank you for all the support and advice you have provided throughout my academic career, I greatly appreciate it. I also want to thank my co-supervisor, Ismail Gogenür, for providing knowledge and insight into the clinical side of my research. Thank you to the rest of the team at Køge. Our collaboration has added new and interesting aspects to my research.

Thank you to all of the members of the Gastro group, both previous and present. I will miss our morning meetings as well as our scientific (and often not so scientific) discussions. Also thank you to my all my colleagues at Roskilde University for the lunch talks and coffee breaks.

I also extend my gratitude to Kevin Myant for allowing me to visit his organoid research laboratory in Edinburgh. Thanks to Sebastian Pohl and Patrizia Cammareri for showing me the ropes in the lab.

Last but not least, I would like to thank my family. Sebastian and August, without your love and support this would not have been possible ♥

Abstract

Colon cancer is one of the most commonly diagnosed forms of cancer, responsible for approximately 1 in 10 cancer-related deaths worldwide. The molecular mechanisms behind the development of colon cancer are heterogeneous but nonetheless important in relation to patient treatment and prognosis. Dysregulation of the intestine specific transcription factor caudal-related homeobox transcription factor 2 (CDX2) has been linked to the development and progression of colon cancer. It plays a critical role in the homeostasis of the heathy colon epithelium, however, the role of CDX2 in colon cancer development is yet to be determined. Evidence suggests that loss of CDX2 expression in colon cancer tumors is linked to poor prognosis, but at the same time CDX2 is regarded as a lineage survival oncogene in colon cancer cell lines. This thesis explores the role of CDX2 in regulation of gene expression in colon cancer cells and what effect CDX2 dysregulation may have on the development and progression of colon cancer.

In **paper I**, a cellular model to investigate effect of a gene of interest was generated by introducing doxycycline inducible CDX2 expression in the LS174T colon cancer cell line. Using this model which allows tight control of CDX2 expression, novel target genes of CDX2 important in colon homeostasis are identified.

The role of CDX2 in regulation of expression of the recurrent colon cancer fusion gene, *VTI1A-TCF7L2*, was investigated in **paper II**. Here we discover that the fusion protein is a dominant negative regulator of Wnt signaling and that it is transcriptionally regulated by CDX2, possibly resulting in aberrant expression of a dominant negative version of TCF7L2.

The surgical stress response is regarded as a risk factor of the recurrence of colon cancer. **Paper III** investigates the effect of perioperative serum from patients undergoing colon cancer surgery on the adhesion abilities of colon cancer cell lines, as well as the effect of CDX2 expression on adhesion. Colon cancer cells shows increased adhesion abilities

in postoperative serum compared to preoperative serum, and the increased adhesion in the postoperative serum is seen to depend on CDX2 expression.

Together, the results in this thesis aim to elucidate the role of CDX2 both in colon homeostasis but also in the development and progression of colon cancer. Novel CDX2 target genes are identified, including the colon cancer fusion protein VTI1A-TCF7L2 possibly linked to the development of colon cancer. CDX2 is shown to affect the adhesion abilities of cultured colon cancer cells, revealing a probable mechanism in which dysregulation of CDX2 may lead to cancer metastasis.

Resumé (Danish)

Koloncancer er en af de hyppigst diagnosticerede former for cancer og ansvarlig for omkring hver 10. cancerrelaterede dødsfald. De molekylære mekanismer bag udviklingen af koloncancer er heterogene men ikke desto mindre vigtige i forhold til patientbehandling og prognose. Udviklingen og progressionen af koloncancer er koblet til dysregulering af den intestinale transkriptionsfaktor caudal-related homeobox transcription factor 2 (CDX2), der spiller en kritisk rolle i homeostasen af raskt kolonvæv. CDX2s rolle i koloncancer udvikling er endnu ikke fuldstændig klarlagt, men forskning tyder på, at nedregulering af CDX2-ekspression i koloncancertumorer er associeret med dårligere prognose for patienten. Samtidig betragtes CDX2 som værende essentiel for overlevelsen af koloncancercellelinjer. Denne afhandling undersøger CDX2s rolle i regulering af genekspression i koloncancerceller, samt hvilken effekt dysregulering af CDX2 kan have på udviklingen og progressionen af koloncancer.

I **artikel I** blev en cellulær model til undersøgelse af gener af interesse udviklet ved at integrere doxycyclin inducerbar ekspression af CDX2 i koloncancer cellelinjen LS174T. Denne model, der tillader fuld kontrol af CDX2 ekspression, kan anvendes til identificering af nye CDX2-regulerede gener, der er vigtige for kolon-homeostase.
Artikel II undersøger CDX2s rolle i reguleringen af ekspressionen af koloncancer fusionsgenet, *VTI1A-TCF7L2*. Vi viser, at fusionsgenet er en dominant negativ regulator af Wnt signalering og at fusionsproteinet er transkriptionelt reguleret af CDX2, hvilket muligvis resulterer i unormal ekspression af den dominant negative form af TCF7L2.
Det kirurgiske stressrespons regnes for at være en risikofaktor for tilbagefald af koloncancer. I **artikel III** undersøges effekten af perioperativt serum fra patienter, der undergår operation for koloncancer, på adhæsionsevnerne af koloncancercellelinjer, og ligeledes effekten af CDX2-ekspression på adhæsionen. Koloncancercellerne viste øget adhæsionsevne i postoperativt serum sammenlignet med preoperativt serum, og den øgede adhæsion i postoperativt serum var afhængig af CDX2 ekspression.

Sammenholdt forsøger resultaterne i denne afhandling at belyse, hvilken rolle CDX2 spiller i både kolon-homeostase samt i udvikling og progression af koloncancer. Artiklerne identificerer nye CDX2-targetgener samt en potentiel kobling mellem koloncancer fusionsproteinet VTI1A-TCF7L2 og udviklingen af koloncancer. Ydermere belyser CDX2s påviste effekt på adhæsionensevnerne hos koloncancercellelinjer en potentiel mekanisme, hvorigennem dysregulering af CDX2 kan føre til cancermetastase.

List of publications

This PhD thesis is based on the following published papers which will be referred to by Roman numerals (I-III) in the text.

Paper I Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression

Pinto R, Hansen L, Hintze J, Almeida R, Larsen S, Coskun M, **Davidsen J**, Mitchelmore C, David L, Troelsen JT, and Bennett EP.
Nucleic Acids Research 2017, 45(13)

Paper II The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2

Davidsen J, Larsen S, Coskun M, Gögenur I, Dahlgaard K, Bennett EP, and Troelsen JT.
PLoS One 2018, 13(7)

Paper III CDX2 expression and perioperative patient serum affects the adhesion properties of cultured colon cancer cells

Davidsen J, Jessen SB, Watt SK, Larsen S, Dahlgaard K, Kirkegaard T, Gögenur I, and Troelsen JT.
BMC Cancer 2020, 20(426)

Additional work performed during the PhD period

HNF4α and CDX2 regulate Intestinal *YAP1* promoter activity

Larsen S, **Davidsen J**, Dahlgaard K, Pedersen OB, and Troelsen JT.

International Journal of Molecular Sciences 2019, 20(2981)

NF-κB signaling stimulated by patient sera in colorectal cancer cell is decreased during the perioperative period

Furbo S, Watt SK, Jessen SB, **Davidsen J**, Sørensen AE, Kirkegaard T, Gögenur I, and Troelsen JT.

Submitted to Journal of Gastrointestinal Cancer

CDX2 regulates interleukin-33 gene expression in intestinal epithelial cell lines

Larsen S, Seidelin JB, **Davidsen J**, Dahlgaard K, Nielsen CH, Bennett EP, Pedersen OB, Coskun M, and Troelsen JT.

Submitted to FEBS letters

Abbreviations

In accordance with guidelines for gene nomenclature, gene names are written in italics while proteins are in non-italics. Human proteins are written in all capital letters, while mouse proteins are written with only the first letter capitalized.

APC	Adenomatous polyposis coli
BMP	Bone morphogenetic protein
Cdk2	Cycline dependent kinase 2
CDX2	Caudal-related homeobox transcription factor 2
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CKIα	Casein kinase Iα
DFS	Disease-free survival
ECM	Extracellular matrix
EGF	Epidermal growth factor
EMT	Epithelial-to-mesenchyme transition
FZD	Frizzled
GATA	GATA binding protein
GPA33	Glycoprotein A33
GSK3β	Glycogen synthase kinase 3β
HEPH	Hephaestin
HNF4α	Hepatocyte nuclear factor 4 α
LAMC2	Laminin subunit γ2
LRP	LDL receptor protein
MAPK	Mitogen-activated protein kinase
MEP1A	Meprin 1A
MET	mesenchyme-to-epithelial transition
MMR	Mismatch repair

MSI	Microsatellite instability
NF-κB	Nuclear factor κB
OS	Overall survival
PFS	Progression-free survival
SSR	Surgical stress response
TCF/LEF	T-cell factor/lymphoid-binding factor
TCF7L2	Transcription factor 7-like 2
VTI1A	Vps-ten-interacting-1a
Wnt	Wingless-type
YAP1	Yes-associated protein 1

Table of Contents

1. Introduction

Colon cancer is the third most commonly diagnosed form of cancer worldwide and is accountable for approximately 1 in 10 cancer-related deaths (Bray et al., 2018). It tends to be slow growing, typically developing over 10-15 years (Dekker et al., 2019). The mortality rate is largely dependent on disease stage at time of diagnosis, but up to 90% are diagnosed at advanced stages resulting in significantly worse survival (Mattiuzzi et al., 2019). The incidence of colon cancer is 3-fold higher in developed compared to developing countries, but with limitations in treatment in developing countries there is less variation in mortality rate (Bray et al., 2018).

To determine patient treatment and prognosis, cancer tumors are currently classified according to a wide range of tumor cell characteristic, including degree of invasion, grade of differentiation, deregulated molecular pathways, and mutational status, as well as lymphocyte density at the tumor center and invasive margin (Angell et al., 2020). The molecular mechanisms underlying the development of colon cancer are heterogeneous but the pathogenesis is nonetheless clinically important, as it is often related to response to treatment and prognosis of the patient (Melo et al., 2013; Sadanandam et al., 2013).

In the last 10 years, dysregulation of the intestine specific transcription factor caudal-related homeobox transcription factor 2 (CDX2) has been implicated in the development and progression of colon cancer (Baba et al., 2009; Graule et al., 2018). It plays a crucial role in proper embryonic development as well as homeostasis of the adult intestinal epithelium (Gao et al., 2009; Keller et al., 2004; Olsen et al., 2013). CDX2 is specifically expressed in the epithelial cells of the adult intestine and it is used as a biomarker to determine the origin of the primary tumor when unknown (Bruun et al., 2018). Reduced expression of CDX2 in colon cancer tumors has been linked to poor prognosis (Bae et al., 2015; Dalerba et al., 2016), and it has been characterized as a tumor suppressor (Mallo et al., 1997). However, in other cancers CDX2 is considered to be an oncogene and it is shown to be essential for the viability of a number of colon cancer cell lines

(Natoli et al., 2013; Salari et al., 2012). Therefore, the effect of CDX2 on the development and progression of colon cancer are yet to be determined.

2. Aim of the thesis

The overall aim of this thesis was to investigate the effect of CDX2 expression on the transcriptional activity of genes involved in colon homeostasis and to explore its role in the development and progression of colon cancer.

Specific aims of the thesis:

- Determine the gene regulatory activity of CDX2 in colon cancer cells (**Paper I**)
- Investigate the role of CDX2 in expression of the recurrent colon cancer gene fusion of the *TCF7L2* and *VTI1A* genes (**Paper II**)
- Explore the effect of CDX2 on colon cancer cell line adhesion abilities (**Paper III**)

3. Background

Colon homeostasis

The epithelium of the colon regenerates rapidly, and homeostasis depends on tight control of proliferation, migration, differentiation, and apoptosis. The lumen of the colon is lined with a single sheet of columnar epithelial cells shaped into tubular glands called crypts of Lieberkühn that extend inwards. Each crypt contains between two and three thousand cells and the renewal of the entire colon epithelium stems from these crypts (Boman and Fields, 2013). The crypt is divided into three compartments each with their own distinct function, seen in figure 1.

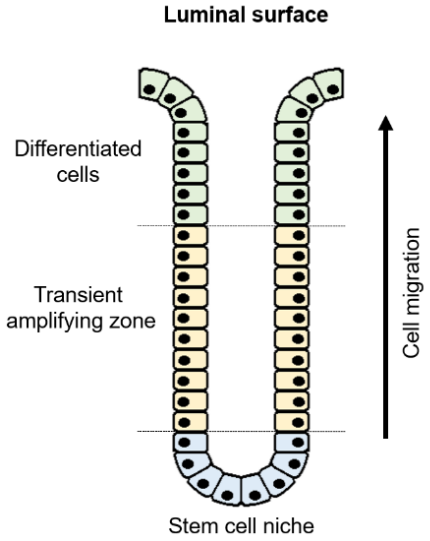


Figure 1: Crypt-villus axis – The intestinal crypts are lined with a single sheet of epithelial cells. The stem cell niche is located at the bottom of the crypts. Proliferating cells migrate up towards the luminal surface though the transient amplifying zone. Proliferation ceases approximately two thirds up the crypt and the cells start to differentiate. Modified from (Davidsen et al., 2018).

The stem cells reside at the base of the crypts in a stem cell niche, and these cells replenish not only their own population but also the cells destined for differentiation. The stem cells continuously divide and produce rapidly proliferating cells called

transient amplifying cells residing in the transient amplifying zone of the crypt. As the cells proliferate, they migrate along the vertical crypt-villus axis towards the luminal surface. When reaching two thirds up the crypt, the proliferation ceases and the cells begin to differentiate (Leedham, 2014). The major types of differentiated cells in the colon include absorptive colonocytes, mucous-secreting goblet cells, and hormone-releasing enteroendocrine cells (Cernat et al., 2014). The differentiated cells reach the luminal surface where they reside until undergoing spontaneous apoptosis with subsequent shedding into the lumen or phagocytosis, about 5-7 days after generation at the bottom of the crypts (Leedham, 2014).

Stem cell maintenance and proliferation

The maintenance of homeostasis of the stem cell niche is crucial in retaining normal colon function. The composition of the stem cell niche is complex and only partly defined but consists of two major components; the extracellular matrix (ECM) consisting of a network of fibrous structural proteins that act as a scaffold and maintains the structure of the colon, and the cellular niche surrounding the stem cell niche. The cellular niche is comprised of pericryptal myofibroblasts, fibroblasts, neuronal cells, endothelial cells, pericytes, and smooth muscle cells that tightly line the basal lamina of the crypt base (Sailaja et al., 2016; Spit et al., 2018). Together, the ECM and the cellular niche provide the crucial signals that regulate intestinal fate. The four main signaling components in stem cell niche maintenance are epidermal growth factor (EGF), Notch, inhibitors of bone morphogenetic protein (BMP), and wingless-type (Wnt) (Spit et al., 2018), seen in figure 2.

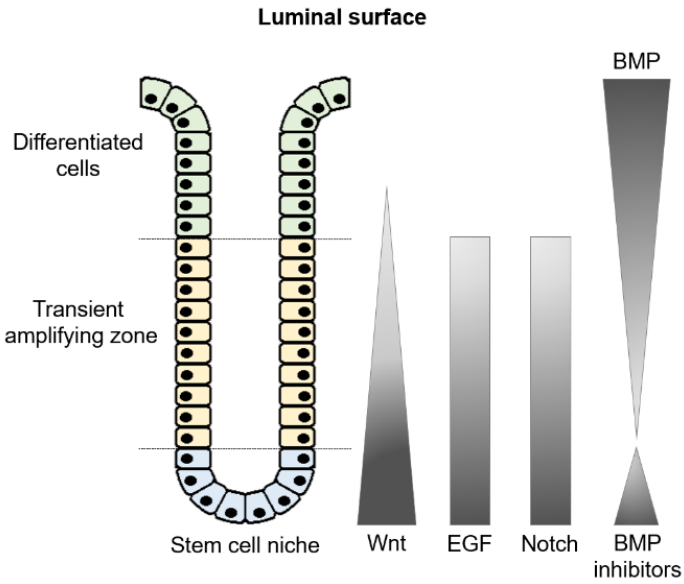


Figure 2: Signaling components of the stem cell niche and proliferation – Wnt, EGF, Notch, and BMP inhibitors are highly expressed at the bottom of the crypts. In the transient amplifying zone Wnt, EGF, and Notch remain highly expressed, while the level of BMP inhibitors decreases and levels of BMP increase towards the luminal surface. As the cells reach differentiation Wnt, EGF, and Notch levels decrease. Modified from (Davidsen et al., 2018).

Activation of EGF receptors activates AKT and ERK pathways that both activate downstream targets involved in regulating cell proliferation and apoptosis. Notch signaling requires direct membrane contact between two adjacent cells in the crypt, one expressing Notch receptors and one expressing Notch ligands. A cell expressing Notch ligands induces Notch signaling in all surrounding cells, which in turn prevents activation of its own pathway. Cells with active Notch signaling enter into the secretory lineage while cells with inactive Notch become absorptive, thereby securing even distribution of secretory and absorptive cells. BMP's counteract the proliferative signals in the intestine and promote cell differentiation. Inhibitors of BMP are present at the base of the intestinal crypts, but expression decreases along the crypt-villus axis (Gehart and Clevers, 2019). As the canonical Wnt signaling pathway has more impact on stem cell maintenance, proliferation and cell fate of intestinal cells than any other pathway, this thesis will primarily focus on this pathway.

The canonical Wnt signaling pathway

Wnt ligands can activate both canonical and non-canonical Wnt signaling pathways. The non-canonical Wnt signaling pathways are involved in cell polarization during epithelial development and intracellular Ca^{2+} influx (Hoppler and Kavanagh, 2007), while the canonical Wnt signaling pathway regulates gene expression, and is known to be the main driver of proliferation of the stem cells in the intestinal crypts (Ong et al., 2014). The canonical Wnt signaling pathway will in the following be referred to as the Wnt signaling pathway.

In the absence of Wnt ligands the Wnt signaling pathway is inactive and the transcriptional co-activator β-catenin is phosphorylated and subsequently degraded in proteasomes. Phosphorylation occurs through a large multi protein degradation complex composed of the kinases glycogen synthase kinase 3β (GSK3β) and casein kinase Iα (CKIα), as well as the scaffold proteins adenomatous polyposis coli (APC) and AXIN (Gehart and Clevers, 2019), as seen in figure 3.

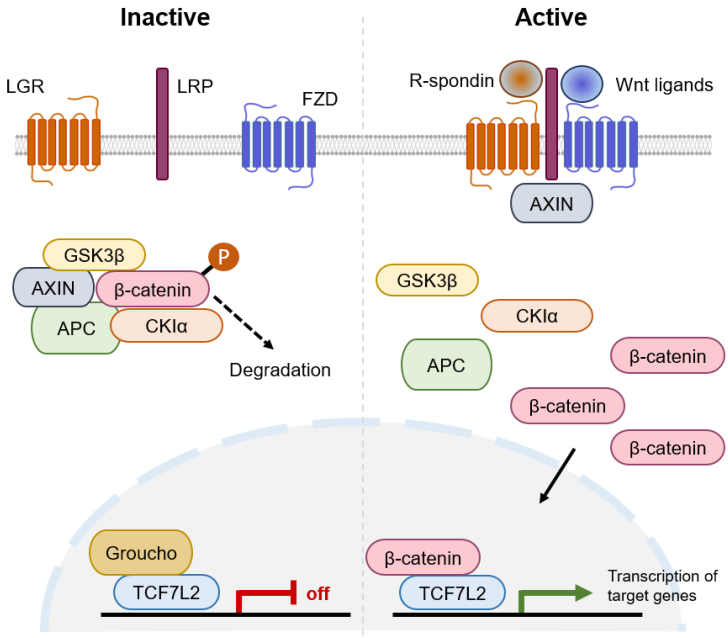


Figure 3: The Wnt signaling pathway – In the absence of Wnt ligand intracellular β-catenin will be degraded though interaction with the degradation complex, comprised of AXIN, APC, CKIα, and GSK3β. Activation of the Wnt signaling pathway occurs when Wnt ligands binds to the FZD and LRP receptor complex. Upon binding, the degradation complex cannot be formed leading to the accumulation of β-catenin, which subsequently enters the nucleus and activates TCF7L2 by displacing Groucho. As R-spondin protects the FZD and LRP from degradation, it presence is required for active Wnt signaling.

The Wnt signaling pathway is activated when secreted Wnt glycoproteins bind to the Frizzled protein (FZD) and low-density lipoprotein receptor-related protein (LRP) transmembrane receptor complex. The activation of the Wnt pathway is also dependent on the presence of R-spondins, secreted and soluble proteins that bind receptors of the LGR family, a stem cell marker (Yan et al., 2017). In the absence of R-spondin, the FZD and LRP receptors are continuously degraded through ubiquitination (Hao et al., 2012). Activation of the Wnt signaling pathway leads to association of AXIN with the LRP receptor and accumulation of the transcriptional co-activator β-catenin in the cytosol. It is subsequently localized to the nucleus where it associates with and activates members of the T-cell factor/lymphoid-binding factor (TCF/LEF) family of

transcription factors, and induces transcription of Wnt target genes. In the colon, the main binding partner of β-catenin is the transcription factor 7-like 2 (TCF7L2) transcription factor, also known as TCF4 (Gehart and Clevers, 2019). In the absence of β-catenin, TCF7L2 forms a complex with the co-repressor Groucho making TCF7L2 act as a transcriptional repressor. Upon activation, β-catenin displaces Groucho turning TCF7L2 into a transcriptional activator (Kriegl et al., 2010; Pinto et al., 2003) that in turn activates transcription of downstream target genes such as *c-myc* and *cyclin D1* (Fevr et al., 2007; Yu et al., 2019).

Wnt signaling has not only been shown to be important in preserving the stem cell niche, but also in maintaining the transient amplifying zone of the crypts (Kuhnert et al., 2004; Pinto et al., 2003). The first hint of the crucial role of Wnt signaling was discovered in a genetically modified mouse model lacking TCF7L2. These mice die within 24 hours after birth and completely lack proliferating compartments in the intestine (Korinek et al., 1998). Proliferation in the crypts of adults remains dependent on Wnt signaling as conditional deletion of both β-catenin and TCF7L2, as well as overexpression of the Wnt inhibitor Dickkopf1, in adult mice both result in a cease in proliferation and loss of intestinal crypts (Fevr et al., 2007; Pinto et al., 2003; van Es et al., 2012). TCF7L2 is not only believed to be important in the proliferative compartments of the crypts, but may play a role as a trigger in the shift from proliferating to differentiating cell. (Kuhnert et al., 2004; Mariadason et al., 2001; Pinto et al., 2003). Conditional loss of TCF7L2 in the colon of adult mice results in enlarged crypts, suggesting continued proliferation and lack of differentiation (Angus-Hill et al., 2011), further underlining the importance of the Wnt signaling pathway in intestinal homeostasis.

Differentiation of the colon epithelium

The transcription factor CDX2 is considered a master regulator of intestinal identity, regulating a number of transcription factors important in colon development and homeostasis. The *CDX2* gene consists of three exons generating a 313 amino acid protein that contains a transactivation domain and a homeodomain for DNA binding. It

binds to gene enhancer and promoter regions on the DNA and can either activate or repress gene expression (Drummond et al., 1997), although results suggest that CDX2 predominantly acts as a transcriptional activator (Pinto et al., 2017) (**Paper I**). In addition to regulating transcription, CDX2 is known to control the active chromatin structure at thousands of sites in the intestine of adult mice, maintaining chromatin access and transcription (Verzi et al., 2013).

The crucial role of CDX2 in differentiation of the colon epithelium has been well established. Knockdown of CDX2 in differentiated Caco-2 cells leads to decreased expression of genes specific for differentiated colon epithelial cells (Natoli et al., 2013). Adult mice with complete knockdown of CDX2 expression develop severe diarrhea and rapid wasting with continued proliferation and lack of differentiation of the colon epithelium (Hryniuk et al., 2012). Organoids derived from mouse small intestine lacking CDX2 expression fail to differentiate but turn into a more gastric like lineage (Simmini et al., 2014), emphasizing the vital role CDX2 plays in differentiation of the colon epithelium.

Regulation of intestinal homeostasis by CDX2

Numerus of genes involved in the function and structure of the colon have been shown to be transcriptionally regulated by CDX2, see figure 4. One of the key functions of the intestine is breakdown and transport of nutrients from the lumen. The iron-transport protein hephaestin (HEPH), which transports dietary iron in the intestinal epithelium, has been shown to be transcriptionally regulated by CDX2. So has the enzyme Meprin 1A (MEP1A), a membrane associated protein that hydrolyses various peptide and protein substrates (Boyd et al. 2010; Coskun et al. 2012; Hinoi et al. 2005). CDX2 is also seen to regulate the expression of lactase, responsible for the hydrolysis of lactose (Troelsen et al., 1997), highlighting its role in breakdown and transport of nutrients.

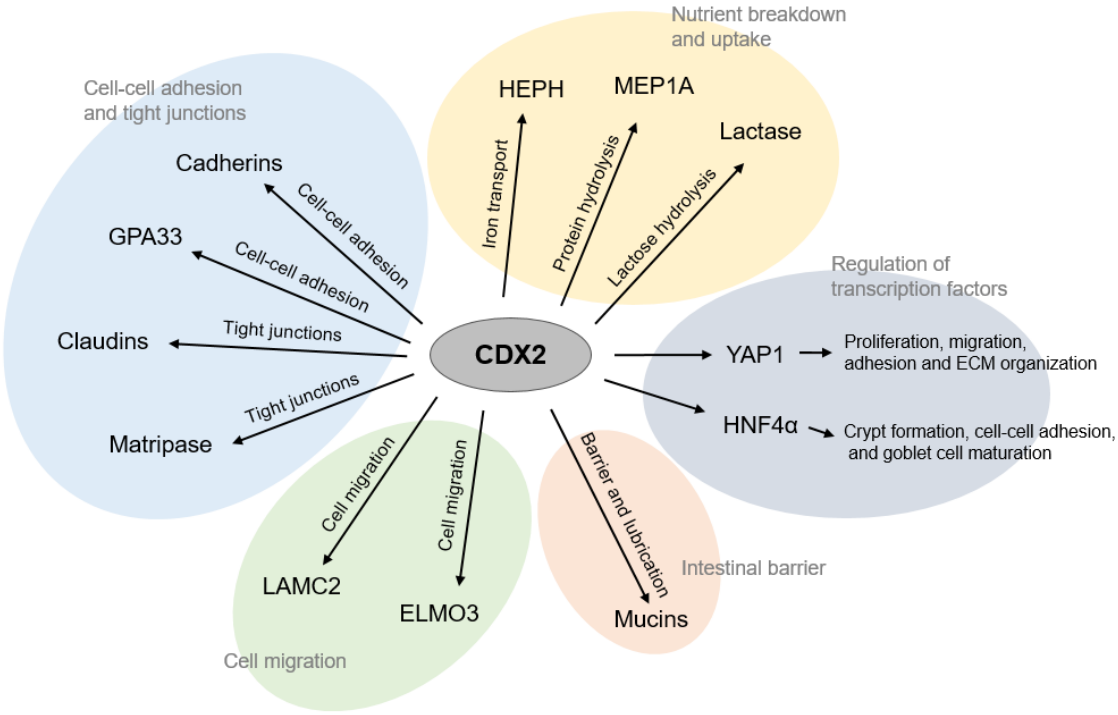


Figure 4: CDX2 and intestinal homeostasis – CDX2 is involved in the transcriptional regulation of many genes important in different aspects of normal intestinal homeostasis, including nutrient uptake and breakdown, regulation of transcription factors, maintaining the intestinal barrier, cell migration, and cell-cell adhesion and tight junctions.

CDX2 is also a key player in upholding the integrity of the intestinal lining by regulating expression of factors important in cell adhesion and tight junctions. It has been linked to members of the cadherin family, molecules that mediate cell-cell adhesion (Hinoi et al., 2002), as well as transcription of glycoprotein A33 (GPA33) (Pinto et al., 2017) (**Paper I**). The functional role of GPA33 remains to be determined. It is expressed in the epithelial cells of the intestine with increased levels at cell-cell junctions (Johnstone et al., 2000), and has been implicated in cell-cell adhesion and maintenance of the intestinal barrier function (Ackerman et al., 2008; Williams et al., 2015). CDX2 is also known to regulate the expression of several claudins, membrane proteins that act as a critical component of the tight junctions in epithelial cells (Bhat et al., 2012; Sakaguchi

et al., 2002; Satake et al., 2008). Further, CDX2 has been shown to regulate the transcription of the *ST14* gene that encodes a type II transmembrane serine protease, also called matriptase (Danielsen et al., 2018). Matripase is involved in maintaining the permeability of tight junctions, possibly by incorporating claudins into intercellular junctions (Buzza et al., 2010). Matripase is under strict post-translation control by two inhibitors encoded by the *SPINT1* and *SPINT2* genes. In addition to regulating transcription of *ST14,* CDX2 also regulates the transcription of the inhibitor encoded by *SPINT1* (Danielsen et al., 2018).

The laminin subunit γ 2 (LAMC2), an ECM protein important in cell migration, is also seen to be transcriptionally regulated by CDX2 (Coskun et al., 2017). As is the expression of the *ELMO3* gene encoding a scaffold protein that stimulates cell migration (Coskun et al., 2010), implicating CDX2 in regulation of cell migration in the colon. A number of mucins have also been shown to be transcriptionally regulated by CDX2, including *MUC5AC*, *MUC5B*, and *MUC6* (Pinto et al., 2017) (**Paper I**). These mucins are secreted by goblet cells and provide a selective barrier to protect the gastrointestinal tract from acid, enzymes, food, and bacteria, as well as creating lubrication (Krishn et al., 2016).

Besides activating the transcription of functional and structural genes, CDX2 is also involved in the transcriptional regulation of other transcription factors important in the colon. In differentiated Caco-2 cells, CDX2 is seen to bind to and activate transcription of the *hepatocyte nuclear factor 4 α (HNF4a)* gene (Boyd et al., 2010), encoding a transcription factor involved in crypt formation, cell-cell adhesion, and maturation of goblet cells (Cattin et al., 2009; Garrison et al., 2006). CDX2 has also been shown to bind a novel enhancer region of the *Yes-associated protein 1* (*YAP1*) gene, a co-transcription factor that regulates many genes important in proliferation, adhesion, migration and ECM organization (Larsen et al., 2019).

Tissue-specific gene expression reflects the coordinated activities of multiple transcription factors, some of which are restricted to individual organs and some that are more broadly expressed. Adult mice with inducible combinatorial deletion of *CDX2*

and either *GATA binding protein 4* (*GATA4)* or *HNF4α* show different crypt consequences. The combination of *CDX2* and *GATA4* deletion results in impaired crypt cell replication, while *CDX2* and *HNF4α* deletion severely compromises cell differentiation but does not affect cell replication. Both combinations produced significantly greater defects and rapid lethality compared to single mutants (San Roman et al., 2015). Thus, CDX2 cooperates with other broadly expressed transcription factors to regulate distinct aspects of intestinal homeostasis.

CDX2 and the Wnt signaling pathway

Evidence suggests that CDX2, in addition to regulating differentiation, also acts as a regulator of the Wnt signaling pathway. In human colon cancer cell lines, the effect of CDX2 on Wnt signaling has been investigated using the TOP/FOP-Flash reporter system (Veeman et al., 2003). In CDX2 knockdown cells, the activity of Wnt signaling is increased while overexpressing CDX2 leads to reduced Wnt signaling (Yu et al., 2019). In several human colon cancer cell lines, CDX2 has been shown to bind directly to β-catenin, thereby inhibiting β-catenin from binding to and activating TCF7L2 (Guo et al., 2010). In mice, it has also been shown that CDX2 regulates the transcription of the protocadherin Mucdhl, which interacts with and inhibits the transcriptional activity of β-catenin, showing that CDX2 also indirectly affects β-catenin activity (Hinkel et al., 2012). In human cancer tissue, expression of CDX2 is seen to be inversely related to β-catenin expression as well as expression of cyclin D1 and c-myc, both downstream target genes of Wnt signaling (Yu et al., 2019) supporting the notion that CDX2 suppresses Wnt signaling in human colon cancer cells.

In addition, the expression of APC, AXIN2, and GSK3β, all members of the β-catenin degradation complex in the Wnt signaling pathway, have also been associated with CDX2. In colon cancer cells, the expression of APC has been shown to be transcriptionally regulated by CDX2 (Olsen et al., 2013). However, in CDX2 knockout and CDX2 overexpression colon cancer cell lines, no change in APC expression can be detected (Yu et al., 2019). Tissue from colon cancer liver metastasis show loss of CDX2

to be associated with nuclear APC levels, but no association is seen with cytoplasmic expression of APC (Tóth et al., 2018). Therefore, the true effect of CDX2 on APC expression remains to be determined.

ChIP-seq data from colon cancer cells show increased CDX2 binding of the AXIN2 promoter region in differentiating cells compared to proliferating, indicating that CDX2 might upregulate the expression of AXIN2 during differentiation (Boyd et al., 2010). AXIN2 has also been shown to transcriptionally regulated by CDX2 in colon cancer cell lines (Olsen et al., 2013). Knockdown of CDX2 in colon cancer cell lines leads to a decrease in expression level of AXIN2, while overexpression of CDX2 increases the AXIN2 mRNA level (Yu et al., 2019), clearly indicating that CDX2 regulates the transcription of AXIN2.

In colon cancer cells, CDX2 has been shown to bind to the promoter region of GSK3β but not activate transcription (Olsen et al., 2013). However, downregulation of CDX2 by TNF-α in colon cancer cells suppresses GSK3β expression (Coskun et al., 2014). Furthermore, knockdown of CDX2 in a colon cancer cell lines leads to a decrease in expression level of GSK3β, while overexpression of CDX2 increases the mRNA level (Yu et al., 2019). These results indicate that while CDX2 binds to the GSK3β promoter, activation of transcription may be dependent on other factors besides CDX2. Despite the conflicting evidence regarding the role of CDX2 in the transcription of APC and GSK3β, it is clear that CDX2 regulates the transcription of AXIN2 and plays a part in regulation of the degradation complex.

Reducing the level of CDX2 in differentiated colon cancer cells results in an increase in transcription of TCF7L2, the main binding partner of β-catenin in the colon. (Boyd et al., 2010). Further, CDX2 has been shown to bind and activate transcription of *HBP1*, a transcriptional repressor that has been shown to suppress Wnt signaling by inhibiting TCF transcription factors (Boyd et al., 2010). These results indicate that CDX2 represses the transcriptional activity of TCF7L2, thereby acting as a repressor of Wnt signaling.

Despite evidence of the inhibitory effect of CDX2 on Wnt signaling through inhibition of β-catenin, activation of transcription of members of the degradation complex and repression of transcription of TCF7L2, other studies indicate a possible interplay between the two. In colon cancer cell lines, a synergistic effect of CDX2 and β-catenin can be observed in activation of the transcription of claudin-1 (Bhat et al., 2012). Further, CDX2 has been shown to occupy some of the same regulatory regions as TCF7L2, resulting in an increase in intestine-specific transcripts compared to TCF7L2 binding alone (Verzi et al., 2010), suggesting that CDX2 may cooperate with Wnt signaling and not only acts as a negative regulator.

Colon cancer development

The median age of diagnosis with colon cancer is approximately 70 years (Brenner et al., 2014). Age is the single most dominant risk factor of colon cancer development, but other identified risk factors include inflammatory bowel disease (Jess et al., 2012), smoking (Liang et al., 2009), obesity (Ma et al., 2013), excessive alcohol consumption (Fedirko et al., 2011), diabetes (Jiang et al., 2011), high consumption of red and processed meats (Chan et al., 2011), and family history of colon cancer (Taylor et al., 2010). While few under the age of 50 are diagnosed, increasing incidence in young adults has been observed, possibly linked to alterations in diet, sedentary lifestyle, and rise of obesity (Cai et al., 2019; Loomans-Kropp and Umar, 2019).

Degree of disease if determined according to TNM classification, describing the extent of invasion of the primary tumor (T), involvement of regional lymph nodes (N), and metastasis to distant sites (M). The three categories are combined into staging groups I-IV, with stage I tumors typically confined to the submucosa or muscular layer of the colon and stage IV presenting with distant metastasis (Piñeros et al., 2019), see table 1. The 5-year survival of patients with colon cancer is 90% in patients with localized cancer (stages I and II), 71% in patients with regional metastasis (stages II and III), and

15% in patients with distant metastasis (stage IV). Thus, the mortality rate is largely dependent on stage of disease at time of diagnosis, and up to 90% of colon cancer patients are diagnosed with advanced stage of disease, typically after onset of symptoms. Advances in screening methods have slightly reduced the number of diagnoses at late stage disease (Mattiuzzi et al., 2019). The primary treatment of colon cancer patients is surgical resection of the tumor as well as adjuvant therapy for patients with stage III and resectable stage IV disease. Some subsets of patients with stage II also benefit from adjuvant therapy, although up to 80% are cured by surgery alone (Kannarkatt et al., 2017). The risk of recurrence after curative resection of colon cancer is up to 30 % (van der Bij et al., 2009). Colon cancer surgery induces a surgical stress response (SSR) that results in reduced anti-tumoral defense as well as an increase in inflammatory factors (Neeman and Ben-Eliyahu, 2013), leading to an increased risk of postoperative recurrence (Heaney and Buggy, 2012).

Table 1: Overview of colon cancer disease stage – Colon cancer is categorized according to degree of invasion of the primary tumor (T), involvement of regional lymph nodes (N), and metastasis to distant sites (M). The combination of the three factors confers to disease stage I-IV, with stage I being early stage and stage IV regarded as late stage.

	Disease stage			
	Stage I	**Stage II**	**Stage III**	**Stage IV**
T	Tumor invades submucosa or infiltrates muscle layer	Tumor invades muscle layer or through peritoneum	Any degree of tumor invasion	Any degree of tumor invasion
N	No metastasis to regional lymph nodes	No metastasis to regional lymph nodes	Metastasis to at least one regional lymph node	Any degree of metastasis to regional lymph nodes
M	No distant metastasis	No distant metastasis	No distant metastasis	Distant metastasis

While well-defined inherited forms account for approximately 2-5% of colon cancer cases, the majority are sporadic. For sporadic colon cancer to develop multiple genetic events are required, and genomic and epigenomic instability is recognized as a hallmark feature of colon oncogenesis (Hanahan and Weinberg, 2011). Several distinct pathways of genomic and epigenomic instability in colon cancer have been described, including chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) (Pino and Chung, 2010). CIN is observed in 70-85% of all sporadic cancers and is characterized by duplication or deletion of either whole chromosomes or parts of chromosomes. MSI tumors are characterized by a defect in the DNA mismatch repair (MMR) system leading to instability in stretches of DNA microsatellites. The CIMP pathway exhibits gene silencing due to hypermethylation of CpG islands. As the definitions of these three pathways are not mutually exclusive, a tumour can exhibit features of multiple pathways (Grady and Markowitz, 2016).

The majority of colon cancers are adenocarcinomas that stem from adenomatous precursor lesions. Over time these lesions grow in size and degree of dysplasia, eventually developing into tumors (Grady and Markowitz, 2016). In sporadic colon cancers that follow the adenoma-carcinoma pathway, inactivation of the *APC* gene is regarded as the early and perhaps initiating event in the multistep process of cancer formation. This loss of action mutation is followed by mutations of essential oncogenes and tumor suppressors (Fearon, 2011). While the majority of colon cancers follow the so-called adenoma-carcinoma pathway, 15-30% arise from early neoplastic serrated lesions histologically characterized by saw-tooth morphology of the epithelial glands in the precursor lesions as well as poor prognosis (Melo et al., 2013; Sakamoto et al., 2017).

The role of CDX2 in cancer development

In early studies, loss of CDX2 expression has been implicated as a common event in tumorigenesis labelling CDX2 a tumor suppressor (Mallo et al., 1997). Mice embryos with complete *Cdx2* knockdown die before birth at embryonic day 3.5, whereas *Cdx2*

heterozygote mice are viable but multiple adenomatous polyps can be observed in the colon between week 12 and 28 (Chawengsaksophak et al., 1997). Mosaic deletion of *Cdx2* in the colon of adult mice results in gastric-like lesions that do not spontaneously evolve into cancer (Balbinot et al., 2018), however, when paired with mice with mutant *Apc* a significant increase in tumor formation can be observed (Balbinot et al., 2018; Hryniuk et al., 2014). By tracking the Cdx2 depleted cells it has been shown that the tumors do not originate from the Cdx2 depleted cells but rather from Cdx2 intact cells (Balbinot et al., 2018), thereby indicating that the Cdx2 depleted cells do not become tumorigenic, but rather stimulate tumorigenesis in adjacent Cdx2 intact *Apc* mutated cells.

Despite the evidence of its tumor suppressor function, there is also evidence that CDX2 may have oncogenic properties. In leukemia and gastrointestinal cancers of the stomach and esophagus CDX2 is considered to be an oncogene (Barros et al., 2012; Rawat et al., 2012; Tamagawa et al., 2012), and studies also indicate oncogenic properties in colon cancer. CDX2 has been shown to be essential for the viability of a number of colon cancer cell lines (Natoli et al., 2013; Salari et al., 2012), and in human colorectal cancers that exhibit chromosomal instability the CDX2 locus has been shown to be amplified (Subtil et al., 2007), indicating that CDX2 plays a role in colon cancer cell survival. Knockdown of CDX2 in colon cancer cells promotes cell proliferation *in vitro* and accelerates tumor formation *in vivo*, while the overexpression of CDX2 suppresses cell proliferation and viability, suggesting that CDX2 inhibits proliferation in colon cancer cells (Yu et al., 2019).

Tumors with a strong reduction of CDX2 primarily belong to the serrated subtype, typically associated with MSI and CIMP, and poor prognosis (Balbinot et al., 2018; Melo et al., 2013). A synergistic oncogenic effect has been observed between loss of CDX2 expression and mutations in the *BRAF* oncogene in serrated tumors, possibly driving the development and progression of colon cancer. In adult mice the simultaneous inactivation of the *Cdx2* gene and mutation of *Braf* dramatically decreased survival compared to each defect on its own, and the mice developed large tumors with

serrated histological features and altered gene expression (Sakamoto et al., 2017; Tong et al., 2017). Correlation between loss of CDX2 expression and *BRAF* mutation has also been observed in human serrated cancers (Landau et al., 2014; Neumann et al., 2018; Olsen et al., 2014; Sakamoto et al., 2017), and a study has implicated *BRAF* in the regulation of CDX2 (Herr et al., 2015), further underlining an interaction between CDX2 and BRAF in early colon cancer development.

CDX2 has been seen to regulate the transcription of several genes that then dysregulated become tumorigenic. The *ST14* gene, encoding matripase, as well as one of its two inhibitors, are both regulated by CDX2 (Danielsen et al., 2018). Matripase has been shown to be dysregulated in a number of cancer types, including cancer of the breast, ovary, uterus, prostate, skin, cervix, and colon (reviwed by Tanabe and List 2017). However, mRNA levels of matripase have also been described to be downregulated in colon cancers, but this downregulation was followed by a downregulation of mRNA of the inhibitors, resulting in the same ratio of matripase to inhibitor (Vogel et al., 2006). The ratio between matripase and its two inhibitors is seen to be increased in cancers, suggesting that this balance is vital in avoiding cancerous development (Parr et al., 2012). As CDX2 regulates both matripase and one of its inhibitors, the dysregulation of CDX2 may be critical in the matripase to inhibitor ratio balance.

Further, CDX2 is shown to bind to enhancer regions of the *YAP1* gene, encoding the co-transcription factor YAP1 (Larsen et al., 2019). It is known to be a key regulator of proliferation, adhesion, migration and ECM organization, has also been linked to cancer progression (Liang et al., 2014; Steinhardt et al., 2008; Tschaharganeh et al., 2013). Nuclear localization and overexpression of YAP1 has been associated with poor prognosis in colon cancer, as well as other types of cancer (Xia et al., 2018; Zygulska et al., 2017).

Alterations in mucin expression, organization, and glycosylation has shown to be associated with colon cancer (Krishn et al., 2016), as well as various other cancers, including breast, ovarian, gallbladder, pancreatic, lung, gastric, and prostate cancer (reviewed by Li et al. 2019). A number of mucins have been identified as CDX2 targets,

including *MUC5AC* (Pinto et al., 2017) (**Paper I**), which is not expressed in healthy colon epithelium. However, its expression has been observed to be upregulated in early stage colon cancers but decrease as dysplasia increases. Patients with MUC5CA-negative colon cancers have worse overall survival and more aggressive disease (Imai et al., 2013; Wang et al., 2017). The dysregulation of CDX2 expression in colon cancers may be linked to the aberrant expression of mucins, possibly driving colon cancer development.

The expression of the tight junction protein claudin-1 is regulated through an interplay between CDX2 and Wnt signaling (Bhat et al., 2012), and increased expression of claudin-1 has been reported in colon cancer (Dhawan et al., 2005; Kinugasa et al., 2010). Mice with mutated *Apc* and claudin-1 expression have significant increase in colon tumor growth (Pope et al., 2014), and colon cancer cells overexpressing claudin-1 injected into nude mice form tumors at a significantly faster rate compared to injecting with control colon cancer cells (Dhawan et al., 2005). In rectal cancer, however, decreased expression of claudin-1 in cancer tissue has been associated with recurrence and poor prognosis (Yoshida et al., 2011), and perhaps the dysregulation of claudin-1 by CDX2, whether it be up- or down-regulated, may be instrumental in cancer development.

A recurrent fusion between the v-SNARE encoded by *Vps-ten-interacting-1a (VTI1A)* and the *TCF7L2* gene has been observed in approximately 3% of primary colon cancer tumors (Bass et al., 2011). Gene translocations and deletions are well known to be associated with cancer development, and the first fusion gene was discovered in chronic myelogenous leukemia (Parker and Zhang, 2013). Other recurrent fusion genes have been observed in other types of cancer, including prostate and lung (Soda et al., 2007; Yu et al., 2010). The fusion between the *VTI1A* and *TCF7L2* genes has been shown to result in a dominant negative form of the TCF7L2 protein, transcriptionally activated by CDX2 (Davidsen et al., 2018) (**Paper II**), indicating that not only does the function of the TCF7L2 protein change but also the expression pattern.

The true role of CDX2 as either a tumor suppressor or an oncogene remains to be determined, but the dysregulation of CDX2, whether it be up or down regulated, seems to play a vital part in the development of colon cancer.

Regulation of CDX2 expression

The expression of CDX2 is regulated, in part, by CDX2 itself in an auto regulatory loop where it activates transcription from its own promoter (Barros et al., 2011; Boyd et al., 2010), as seen in figure 5. A combination of different factors have also been shown to regulate the intestine specific expression of CDX2, including TCF7L2, HNF4α, GATA6, and nuclear factor κB (NF-κB) (Benahmed et al., 2008; Boyd et al., 2009; Kim et al., 2002). Further, methylation at two distinct promoter regions of the *CDX2* gene correlates with diffuse, partial, or absent CDX2 expression in colon cancer cell lines (Graule et al., 2018).

CDX2 protein can be detected all along the crypt-villus axis but its activity is regulated by post-translational phosphorylation of different sites. Phosphorylation at serine 60 via the mitogen-activated protein kinase (MAPK) pathway is observed in the proliferating compartment of the intestinal crypt, resulting in decreased transactivation capacity (Rings et al., 2001). Interestingly, when CDX2 is phosphorylated by the MAPK family member p38 the transactivation capacity increases (Houde et al., 2001). When phosphorylated by cycline-dependent kinase 2 (Cdk2), CDX2 is marked for rapid degradation by the ubiquitin/proteasome pathway (Boulanger et al., 2005; Gross et al., 2005).

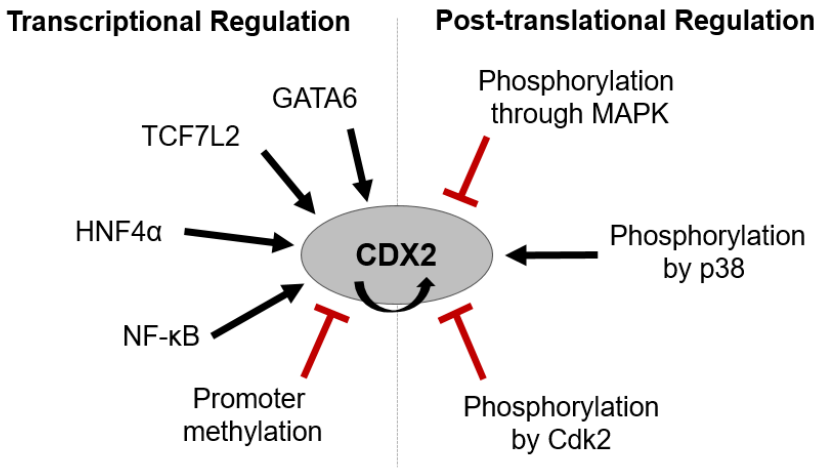


Figure 5: Regulation of CDX2 expression – The GATA6, TCF7L2, HNF4α and NF-κB transcription factors regulate the transcription of CDX2. Methylation of the promoter region of CDX2 results in inhibition of transcription. The activity of CDX2 is regulated by phosphorylation. When phosphorylated through the MAPK and by Cdk2, activity of CDX2 is repressed, while phosphorylation by p38 activates CDX2.

CDX2 as a prognostic marker

As CDX2 expression is retained in most colon cancers it is commonly used as a tissue biomarker to determine the origin of primary tumor. However, a subgroup of patients present with tumors with loss or reduction of CDX2 expression (Baba et al., 2009; Hinoi et al., 2001). The *CDX2* gene is rarely mutated (Xia et al., 2009) but epigenetic changes, such as hypermethylation of the *CDX2* promoter, are believed to lead to transcriptional downregulation or loss of expression of CDX2 (Dawson et al., 2014; Graule et al., 2018). Loss of CDX2 expression is seen as a signature for undifferentiated colon cancer cells and poorly differentiated cancers are seen to be more aggressive with a higher risk of recurrence (Merlos-Suárez et al., 2011). Therefore, the possible prognostic feature of CDX2 status has been thoroughly investigated but despite the substantial amount of publications, the applicability of CDX2 as a prognostic factor is debatable.

The majority of studies investigating CDX2 as a prognostic marker determine an association with loss of CDX2 with either worse overall survival (OS), progression-free survival (PFS), and/or disease-free survival (DFS). Some, however, find no independent prognostic value in loss of CDX2 expression, or find that prognostic value is limited to certain cancer stages or molecular and phenotypical subtypes, as reviewed in table 2.

Table 2: CDX2 as a prognostic marker – Overview of literature that investigates the applicability of CDX2 status as a prognostic marker for prognosis in colon cancer patients. The table shows surname of fist author, year of publication, stage of disease, and whether an association between loss of CDX2 expression and prognosis was observed.

First author	Year	Stage	Loss of CDX2 expression
Asgari-Karchekani et al.	2019	I-IV	No impact
Bae et al.	2014	III IV	Associated with lower PFS. Associated with lower PFS and OS.
Bonetti et al.	2017	I-III	Associate with lower OS and DFS in tumors with poorly differentiated clusters of cells.
Bruun et al.	2018	IV	Associated with lower 5-year DFS
Cecchini et al.	2019	II	No impact
Dalerba et al.	2016	II & III	Associated with lower 5-year DFS.
Dawson et al.	2014	I-IV	No impact
Graule et al.	2018	I-IV	Associated with lower 5-year DFS
Hansen et al.	2018	II	Associated with lower 7-year DFS
Ma et al.	2019	II & III	Associated with lower DFS in MMR deficient tumors
Neumann et al.	2018	I-IV	Associated with lower OS
Nishiuchi et al.	2019	II & III	Associated with lower 5-year OS and DFS
Olsen et al.	2016	I-IV	No impact on risk of recurrence.
Pilati et al.	2017	II & III	Associated with lower OS and DFS
Ryan et al.	2018	I-IV	No impact in MMR deficient tumors
Shigematsu et al.	2018	III & IV	No impact in patients with liver metastasis

Slik et al.	2019	II	Associated with lower 5-year DFS in MSI tumors
Xu et al.	2019	I-IV	Associated with lower OS and DFS
Zhang et al.	2017	IV	Associated with lower OS and PFS in metastatic colon cancers

DFS: disease-free survival. MMR: mismatch repair. MSI: microsatellite instability. OS: overall survival PFS: progression-free survival.

The difference in the prognostic value of CDX2 on survival may be attributed to multiple factors, including the method used to determine CDX2 status and well as the cut-off value for a negative CDX2 status. Assessing CDX2 status from whole-slide immunohistochemistry is often difficult, as there may be areas with CDX2 loss despite most of the tumor being positive (Cecchini et al., 2019). The disease stage of the patients included also varies. Some studies include patients with stage I to stage IV disease, while others investigate patients within one stage or even within one distinct molecular subgroup. The number of patients and time to follow-up also varies, and when taking all these factors into account, the discrepancy of the results may not be surprising.

While CDX2 is regarded as an independent prognostic factor in some studies, others speculate that other molecular features of the tumor should also be taken into account. Mutation of the *BRAF* oncogene have been associated with loss of CDX2 expression (Dawson et al., 2014; Neumann et al., 2018; Slik et al., 2019), and CDX2 loss has displayed strong prognostic value in patients with *BRAF* mutated tumors (Bruun et al., 2018).

Tumors with MMR deficiency have been associated with loss of CDX2 expression (Graule et al., 2018; Olsen et al., 2016), and the prognostic impact of CDX2 is also seen to be dependent on whether the tumor is MMR deficient or proficient. On one hand, in patients with MMR deficient tumors loss of CDX2 has been shown to have no impact on survival (Ryan et al., 2018), but loss of CDX2 expression also been observed to result in worse disease-free survival (Ma et al., 2019). MMR deficient tumors are often characterized by a favorable prognosis (Hansen et al., 2018), but may in combination

with loss of CDX2 expression become a more aggressive subtype of colon cancer. Pilati et al. also shows that the prognostic abilities of CDX2 depends on the molecular subtype of the cancer, and that it can only predict prognosis in some colon cancer subtypes (Pilati et al., 2017).

In addition to molecular subtypes, CDX2 status may also be paired with other features to more precisely determine prognosis. Loss of CDX2 expression has been seen to have improved prognostic value in combination with tumor stage (Xu et al., 2019), Muc2 expression (Cecchini et al., 2019), SATB2 expression (Ma et al., 2019), and number of poorly differentiated clusters of cancer cells within the tumor (Bonetti et al., 2017). To summarize, the applicability of CDX2 as a prognostic marker remains controversial, despite a majority of studies finding association between loss of CDX2 expression and poor prognosis.

CDX2 expression and metastasis

The molecular mechanism behind the possible prognostic abilities of CDX2 remains to be elucidated, however several steps in cancer progression and metastasis have been linked to CDX2. Tumor budding cells frequently show reduced CDX2 expression, despite the tumor being CDX2 positive (Graule et al., 2018; Hansen et al., 2018), associating downregulation of CDX2 with the metastatic process.

Alternation in the adhesion abilities of cancer cells is believed to be a critical element in cancer cell metastasis (Läubli and Borsig, 2019). In order to disengage from the primary tumor, cancer cells downregulate cell-cell adhesion through modification of the cadherin-catenin complex (Aamodt et al., 2010), as well as downregulate expression of claudins, reducing the integrity of tight junctions (Tabariès and Siegel, 2017). As previously mentioned, CDX2 has been shown to regulate the expression of both cadherins and claudins (Bhat et al., 2012; Hinoi et al., 2002; Sakaguchi et al., 2002; Satake et al., 2008). Furthermore, loss of CDX2 expression has been shown to affect adhesion abilities of colon cancer cells (Davidsen et al., 2020) **(Paper III)**, implicating CDX2 in the downregulation of adhesion in tumor budding cells. As circulating tumor

cells attach to metastatic sites the expression of selectins, integrins, and members of the immunoglobulin superfamily are upregulated (Gallicchio et al., 2008; Robertson et al., 2009; Tremblay et al., 2008). CDX2 has been shown to mediate E-selectin ligand expression in colon cancer cells (Sakuma et al., 2012), thereby also involving CDX2 in cancer cell attachment at distant sites of metastasis.

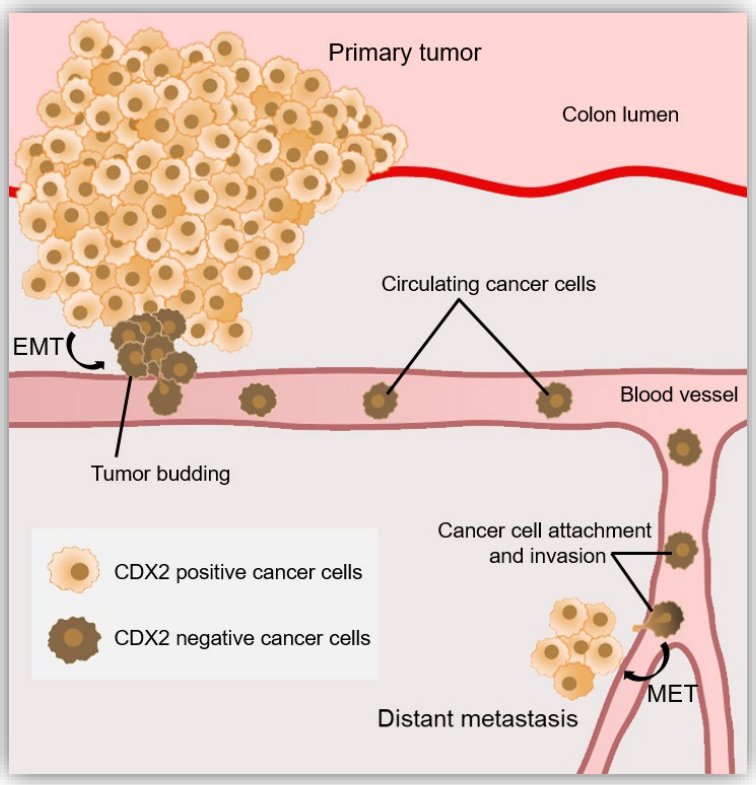


Figure 6: Proposed model of CDX2 expression in metastasizing cancer – For cancer cells to metastasize, the cells must first disengage from the primary tumor. The cells downregulate CDX2 expression leading to EMT and enter circulation. When attaching at distant sites, CDX2 expression is upregulated leading to increased adhesion and MET, resulting in establishment of distant metastasis.

Loss of intestinal identity though the downregulation of CDX2 has been suggested as a precursor to epithelial to mesenchymal transition (EMT) in metastatic colon cancer cells, as seen in figure 6. As the circulating cancer cells establish metastasis, CDX2 expression is reestablished and the cancer cells undergo mesenchymal to epithelial

transition (MET) (Brabletz et al., 2004; Zhang et al., 2015). When overexpressing CDX2 in colon cancer cell lines, a decrease in mobility and dissemination of cancer cells has been observed (Gross et al., 2008; Zheng et al., 2011), further implicating CDX2 in the metastatic process.

4. Publications

This section contains the three published papers included in this thesis.

Paper I

Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression

Pinto R, Hansen L, Hintze J, Almeida R, Larsen S, Coskun M, **Davidsen J**, Mitchelmore C, David L, Troelsen JT, and Bennett EP.
Nucleic Acids Research 2017, 45(13)

Investigating the true function of essential genes can be challenging. The aim of this study was to develop a method where inducible transcriptional elements could be integrated into cultured cells, enabling tight control of gene expression with the use of doxycycline. To demonstrate the possibility of using this method to investigate gene function, the LS174T human colon cancer cell line was genetically engineered as a bi-allelic CDX2 knockout containing doxycycline inducible CDX2 expression.
This paper demonstrates that the PrIITE systems allows for induction of expression of the gene of interest, while at the same time enabling control over protein levels of the gene of interest. This approach revealed novel downstream effector genes of CDX2, indicating the extensive role of CDX2 in gene expression regulation. Further, normal expression of downstream effector genes was highly dependent on the level of CDX2 protein expressed. The method developed in this study can be used to study the functions of other essential genes in a dose-dependent manner.

This paper is a result of collaboration between multiple institutions. My contribution as co-author was determining the transcriptional regulation of promoter reporter constructs by CDX2. Results are presented in figure 6 and supplementary figure 3.

Nucleic Acids Research, 2017 **1**
doi: 10.1093/nar/gkx371

Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression

Rita Pinto[1,2,3,4], Lars Hansen[4], John Hintze[4], Raquel Almeida[1,2,3,5], Sylvester Larsen[6,7], Mehmet Coskun[8,9], Johanne Davidsen[6], Cathy Mitchelmore[6], Leonor David[1,2,3], Jesper Thorvald Troelsen[6] and Eric Paul Bennett[4,*]

[1]i3S, Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal, [2]Ipatimup, Institute of Molecular Pathology and Immunology of the University of Porto, Porto, Portugal, [3]Faculty of Medicine of the University of Porto, Porto, Portugal, [4]Copenhagen Center for Glycomics, Departments of Cellular and Molecular Medicine and Odontology, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark, [5]Department of Biology, Faculty of Sciences of the University of Porto, Porto, Portugal, [6]Department of Science and Environment, Roskilde University, Roskilde, Denmark, [7]Department of Clinical Immunology, Naestved Hospital, Naestved, Denmark, [8]Department of Gastroenterology, Medical Section, Herlev Hospital, University of Copenhagen, Copenhagen, Denmark and [9]The Bioinformatics Centre, Department of Biology & Biotech Research and Innovation Centre (BRIC), University of Copenhagen, Copenhagen, Denmark

Received October 03, 2016; Revised March 30, 2017; Editorial Decision April 10, 2017; Accepted April 27, 2017

ABSTRACT

Tetracycline-based inducible systems provide powerful methods for functional studies where gene expression can be controlled. However, the lack of tight control of the inducible system, leading to leakiness and adverse effects caused by undesirable tetracycline dosage requirements, has proven to be a limitation. Here, we report that the combined use of genome editing tools and last generation Tet-On systems can resolve these issues. Our principle is based on precise integration of inducible transcriptional elements (coined PrIITE) targeted to: (i) exons of an endogenous gene of interest (GOI) and (ii) a safe harbor locus. Using PrIITE cells harboring a GFP reporter or CDX2 transcription factor, we demonstrate discrete inducibility of gene expression with complete abrogation of leakiness. CDX2 PrIITE cells generated by this approach uncovered novel CDX2 downstream effector genes. Our results provide a strategy for characterization of dose-dependent effector functions of essential genes that require absence of endogenous gene expression.

INTRODUCTION

Historically, analysis of the molecular genetic mechanisms underlying cell fate and animal phenotypes has been studied by abrogating gene function in cellular and animal model systems. Initially this has been accomplished by random mutagenesis (1–3), homologous recombination (4) and recently by the use of precise genome editing technologies that allow for target inactivation of any GOI in cells, tissues and animal models (5–7). However, the fact that the genetic lesions induced are static render these approaches inadequate in situations where swift reversal of gene function is desired or in cases where the GOI plays an essential function for cellular survival. Thus, alternative approaches have to be employed in these situations. One commonly used alternative is based on gene 'knock down' by RNAi/shRNA (8). Although the successful application of these technologies in cell lines is well documented, knock down strategies are hampered by lack of quantitative and absolute inactivation of gene function, which makes this approach problematic in situations where downstream gene functional studies require complete gene inactivation. In these situations, the use of inducible gene expression systems has shown to be a powerful methodology that allows for: (i) control of gene expression levels of potentially toxic gene products that could have adverse side effects on cell growth and survival when expressed constitutively, (ii) temporal and spatially controlled activation of genes and proteins and (iii) analysis of cellular gene dose/response effects. Various inducible gene ex-

*To whom correspondence should be addressed. Tel: +45 35326630; Email: epb@sund.ku.dk
Present address: Rita Pinto, Department of Biosciences, Faculty of Mathematics and Natural Sciences, University of Oslo, Blindern, 0316 Oslo, Norway.

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

pression systems have been described in the literature based on cre-lox P system (9), myxovirus resistance 1 promoter (10), estrogen receptor (11), optogenetics (12), ecdysone-induciblity (13) or tetracyclin (Tet)-Off/On systems (14–17). The latter systems are probably the most commonly used inducible systems for which a plurality of reagents have been developed, published and are commercially available (18). However, one disadvantage of Tet-inducible systems and the majority of the aforementioned inducible systems, is their well described 'leakiness' (19–23). Furthermore, inducible expression in both transient and stable expressing cells have indicated that cells respond differently to induction, which has been attributed to heterogeneity in chromosomal integration of the inducible gene elements in individual cells leading to non-homogenous induction responses (21). Importantly, well known side effects of tetracycline (and its derivatives) on cell fitness, in particular after long term treatment, caution for its use in biomedical research when used at traditional concentrations >100 ng/ml (24–26).

Thus, there is a yet unmet need in the field to improve the tightness of the available inducible systems. We reasoned that the observed 'leakiness' with the most commonly used 'Tet-On' system (17,27) is due to the uncontrolled randomness of integration of the genetic elements encoding both the ectopically expressed transactivator and the inducible GOI, leading to muddled inducible gene expression with pleiotropic downstream effects dependent on the activation conditions used. Therefore, we hypothesize that by integrating a defined number of transactivator and inducible transcriptional elements at defined cellular genomic loci we can (i) circumvent the 'leakiness' issue and (ii) lower the Dox concentration needed for induction below the levels causing cellular stress. Thus, by taking advantage of precise genome editing and last generation Tet-On platforms (18,27), the objectives of this study were to establish a flexible 'non-leaky', minimal Dox concentration requiring isogenic knockout-rescue system. We built our cellular model system on the colorectal cell line LS174T and by mono- or bi-allelic targeting of constitutively expressing transactivator (Tet3G) elements (TET3G) to one locus and inducible GFP-reporter elements to another safe harbor locus we demonstrate minimal Dox requirement, no leakiness and reversibility of the system. We next demonstrate the utility of the isogenic PrIITE system by mono- or bi-allelic integration of inducible CDX2 transcription factor elements into PrIITE cells and re-confirm the reversibility and lack of leakiness of the system. We also for the first time demonstrate, that Tet-On leakiness is related to the cellular copy number of integrated transactivator elements and not the number of integrated inducible elements of the inducible system.

Finally we confirm the complete absence in leakiness of the system by RNA-seq and in combination with ChIP-seq identify several novel genes directly transcriptionally controlled by CDX2 (Figure 1). The selection of CDX2 as our target gene was based on its relevance as a key regulator of intestinal differentiation, with many downstream targets that can be assessed as read-outs (28).

MATERIALS AND METHODS

ZFN gene targeting plasmids and plasmid donor construction

CDX2 and AAVS1 CompoZr ZFN plasmids for human *CDX2* and *AAVS1* were obtained from Sigma targeting exon 1 of *CDX2* (AACTTCGTCAGCCCCccgcagTACCC GGACTACGGCGGTT) and intron one of the PPP1R12C gene at the *AAVS1* hotspot AAV integration site (ACCCCA CAGTGGggccacTAGGGACAGGAT), respectively. ZFN binding sites are shown in upper case and linker cut site in lower case lettering. ZFNs were tagged with 2A peptide fused GFP or Crimson as described recently (29). Fluorescent protein tagging allows for FACS enrichment and improves efficiency in obtaining correctly targeted clones.

Donor construct was designed based on the previously described approach by Maresca *et al.* (30) with the only modification that inverted ZFN binding sites were positioned flanking the entire donor insert. A synthetic ObLiGaRe CDX2 donor vector frame work possessing inverted CDX2 ZFN binding sites (AACCGCCG TAGTCCGGGTAccgcagGGGGCTGACGAAGTT) flanking a XhoI/EcoRV linker was generated (EPB64, Genewiz, USA, Addgene ID#90017). A CMV-Tet3G-SV40UTR TET3G transactivator encoding fragment (pTet3G, Clontech/Takara, USA) was XhoI/HindIII excised (HindIII overhang bluntended) and inserted into XhoI/XbaI (overhangs bluntended) site of EPB64-donor vector generating *pCDX2-pCMV-TET3G-ObLiGaRe* donor vector.

Various inducible gene expression ObLiGaRe donor constructs targeted to the *AAVS1* safe harbor locus were generated based on available *AAVS1* safe harbor CompZr ZFN binding element information (Sigma-Aldrich, USA). A *AAVS1* ObLiGaRe donor vector framework was generated (EPB58, Genewiz, USA, Addgene ID#90016) and designated pObLiGaRe-*AAVS1*, possessing the respective inverted ZFN binding elements (5′-atcctgtcccta ggccacccactgtgggt-3′) flanking an EcoRV multiple cloning site linker (see Supplementary Figure S11 for vector design). A full-length codon optimized CDX2 (CDX2opt) construct was generated (EPB40, GeneArt/ThermoFisher, USA and SalI/BamHI insert cloned into SalI/BamHI sites of pTRE3G (Clontech/Takara, USA) generating pTRE3G-CDX2opt.

A BamHI fragment encoding a Golgi targeted EYFP fusion protein, T2EYFP (31) was blunt end cloned into the EcoRV site of pTRE3G generating *p*TRE3G-T2EYFP. T2EYFP encodes the N-terminal Golgi targeting and retention sequence fused to EYFP. The XbaI/XhoI Tet responsive expression fragments from both *p*TRE3G-CDX2opt and *p*TRE3G-T2EYFP were excised and blunt end cloned into the EcoRV site of pObLiGaRe-AAVS1 (EPB58) generating *p*AAVS1-TRE3G-CDX2opt and *p*AAVS1-TRE3G-T2EYFP. For *p*AAVS1-TRE3G-T2EYFP, the donor insert was excised with PacI/PmeI and inserted into PacI/PmeI site of a modified AAVS1 targeting vector (EPB71, Genewiz, USA, Addgene ID#90018) where PacI/PmeI cloning sites are flanked by insulator sequences and the inverted AAVS1 ZFN binding sites. See Supplementary Fig-

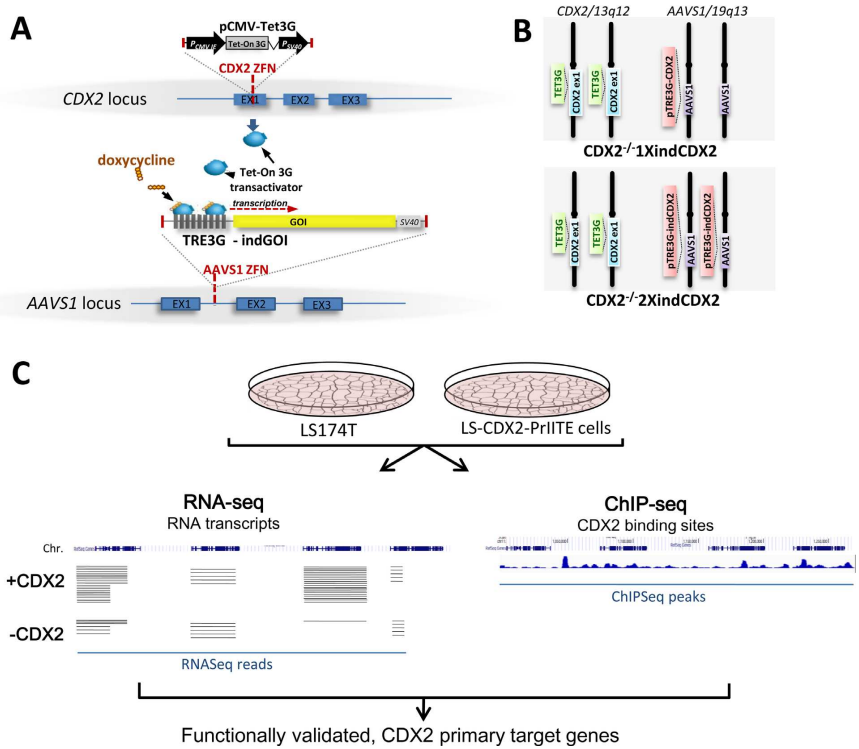


Figure 1. Overview of the established isogenic knockout-rescue model system based on precise integration of inducible elements (PrIITE) in LS174T. (**A**) pCDX2-pCMV-TET3G-ObLiGaRe donor vector (pCMV-Tet3G) was targeted to CDX2 exon1 by ZFN mediated targeted integration followed by pTRE3G-inducible Gene Of Interest (GOI) ObLiGaRe donor vector (pTRE3G-indGOI) targeting to the safe harbor AAVS1 locus. (**B**) Schematic illustration of the isogenic inducible CDX2 LS174T cell model system based on the mono- (upper panel) or bi- (lower panel) allelic precise genome targeting of CDX2 inducible elements, CDX2$^{-/-}$1XindCDX2 and CDX2$^{-/-}$2XindCDX2 respectively. (**C**) Illustration of the experimental workflow including deep transcriptomic analysis (RNA-seq) of wild type LS174T and isogenic indCDX2-PrIITE cells exposed to variable concentrations of doxycycline inducer and CDX2 target validation by genome-wide CDX2 ChIP-seq (ChIP-seq) analysis.

ure S11 for vector constructs used and generated. All plasmids were Sanger sequencing verified.

Cell culture and transfection

LS174T (ATCC-CL-188) human intestinal cell line was maintained in Ham's nutrient mixture F12/Dulbecco's modified Eagle's medium (1:1) supplemented with 10% fetal bovine serum and 1% glutamine. Different target 'knock in' strategies were undertaken, including use of 800 bp homology arms flanking the donor integration cassette D (data not shown). Successful LS174T targeted KI was only obtained when the ObLiGaRe strategy for improved target specific 'knock in' integration of donor constructs was employed (30). In brief, the ObLiGaRe strategy is based on use of existing ZFN binding elements that in an 'inverted' orientation flank the donor DNA construct. By co-nucleofection of ZFN's and donor DNA, double stranded breaks at both the specific chromosomal target site and flanking the donor plasmid will occur allowing cellular repair pathway mediated target specific integration of the linearized donor at the desired target site.

For stable *pCDX2-pCMV-TET3G-ObLiGaRe* integration the first exon of the *CDX2* gene was targeted through CDX2 CompoZr ZFNs driven integration (Sigma-Aldrich, USA) (AACTTCGTCAGCCCCccgcagTACCCGGACTACGGCGGTT, left and right ZFN binding sites capitalized and linker cloning site in lower case). 1×10^6 LS174T cells were transfected by nucleofection, simultaneously with the GFP/E2 Crimson tagged CDX2 ZFN plasmids (2 μg each) and the pCDX2-CMV-TET3G-donor vector (5 μg). Nucleofector solution T (Lonza, CH)/Nucleofector program T-020 were used for the electroporation procedure in an Amaxa Cell Line Nucleofector device (Lonza, CH). Cells were then cultured for 6 h at 37°C to stabilize, and then moved on to a 30°C cold shock. Two days after transfection, cells expressing both GFP and E2 Crimson were sorted out in a FACS ARIA III (BD BioSciences, USA) as previously described (29). The cell bulk was plated out to grow in collagen-coated plates and two rounds of cloning were then employed.

Single cell clones were screened by junction PCR using a primer flanking the 5′ *CDX2* target locus and a reverse oriented primer localized within the integrated *pCMV-TET3G* cassette (Supplementary Figure S1A and B). Based on IDAA assay (32), one of them (5B5) showed the presence of wt allele, representing heterozygozity for *p*CMV-TET3G integration at the *CDX2* locus (clone named $LS^{CDX2+/-}$), while the other two clones (3D6 and 5E2) were found to possess successfully integrated *p*CMV-TET3G transactivator at both alleles representing homozygozity for *p*CMV-TET3G (clones named $LS^{CDX2-/-}$#1 and $LS^{CDX2-/-}$#2 respectively).

1×10^6 cells from a *CDX2* KO clone containing mono or bi-allelic *p*CMV-TET3G transactivator inserted into the two *CDX2* alleles were then transfected by nucleofection with GFP/E2 Crimson tagged AAVS1 ZFN plasmids (2 μg each) and 5 μg *p*AAVS1-TRE3G-CDX2opt or *p*AAVS1-TRE3G-T2EYFP.

Clone characterization by polymerase chain reaction (PCR) and IDAA

In order to identify correct integration of *pCDX2*-CMV-TET3G, a region comprising a part of *CDX2* exon 1 upstream the ZFNs cutting site and left 800 bp homology arm and a part of the CMV promoter was amplified using Expand Long Template PCR System (Roche Applied Science, GE) (5′-CMV-TET3G junction PCR). Similarly, the identification of correctly integrated *p*AAVS1-TRE3G-CDX2opt or *p*AAVS1-TRE3G-T2EYFP fragments into *AAVS1* locus was performed by amplification of a region comprising a part of *AAVS1* upstream the ZFNs cutting site and a part of the *TRE* promoter (5′-TRE3G-CDX2opt or TRE3G-T2EYFP junction PCR). PCR reaction mixture consisted of 100 ng DNA, 2.5 μl buffer 1, 3.0 μl 1.25 mM dNTPs mix, 0.25 μl each primer at a concentration of 25 μM, 0.1 μl enzyme mix and water to a final volume of 25 μl. All junction PCR primers were obtained from TAG Copenhagen A/S, Denmark and are listed in Supplementary Table SIV. Amplification was done using the following touch down protocol. After preheating for 5 min at 95°C, 12 cycles were performed starting with denaturation for 45 s at 95°C, annealing for 15 s at 74°C with a decrease in annealing temperature of –1°C/cycle, and 2 min at 72°C, followed by an additional 25 cycles of 45 s at 95°C, 15 s at 64°C and 2 min at 72°C, followed by a final extension of 3 min at 72°C. PCR products were run in a 1.2% agarose gel, bands were gel purified and sequence confirmed by Sanger sequencing.

FACS-enriched stable clones (29) were screened by junction PCR using primers flanking the junction between the *CDX2* or *AAVS1* genes and the integrated cassette (Supplementary Figures S1 and S2).

Presence of unmodified CDX2 or AAVS1 target (WT allele presence test) was performed using the recently described IDAA method (32) and protocol guidelines (33). In brief, due to size of the donor constructs used (>2 kb, respectively) the respective CDX2 or AAVS1 ZFN target sites can only be successfully amplified if integration has not occurred at the respective target loci. PCR was performed using 100 ng DNA in 25 μl using AmpliTaq Gold (ABI/Life Technologies) for *CDX2* locus or TEMPase Hot Start DNA Polymerase (Ampliqon A/S, DK) for *AAVS1* locus using the recently described IDAA/tri-primer amplification conditions comprising primers flanking the respective target site in combination with a universal 6-FAM 5′-labeled primer (FamF), specific for a 5′-overhang attached to the forward primer. IDAA primers used were purchased from TAG Copenhagen A/S and are listed in Supplementary Table SIV. Fluorescently labeled amplicons were then analyzed by capillary electrophoresis based fragment analysis using an ABI3030 instrument (Applied Biosystems/Life Technologies, USA). Raw data obtained was analyzed using Peak Scanner Software V1.0 (Applied Biosystems/Life Technologies, USA).

Induction with Dox

The Tet-On3G system was induced with Dox (Sigma-Aldrich, USA) in a range between 0.004 and 4 μg/ml for CDX2 knock-in (KI) clones and 0.001–5 μg/ml for T2EYFP KI clones. Dox was added to the medium every 24 h for one or 2 days, and medium was exchanged every 48 h after Dox removal. Controls where Dox was not added to the medium were used (0 μg/ml of Dox).

Western blot of cell lysates

Whole-cell extracts were obtained by resuspension of cell pellets in RIPA buffer (50 mM Tris–HCl pH 7.4, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.1% sodium dodecyl sulphate) in the presence of complete protease inhibitors cocktail (Roche Applied Science, GE). Quantification of total protein was determined by bicinchoninic acid protein assay (Thermo Scientific™ Pierce, USA). 15 μg of protein extracts were then analyzed by standard SDS-PAGE, transferred to a nitrocellulose membrane (GE Healthcare Life Sciences, USA) and blotted on at 4°C with mouse monoclonal primary antibodies CDX2-88 1:500 (Biogenex, USA), anti-TetR 1:1000 (Clontech/Takara, USA), anti-MUC2, undiluted hybridoma supernatant and anti-actin 1:8000 (Santa Cruz Biotechnology, USA) in 5% BSA in PBS. Membranes designed for MUC2 blotting were previously treated with neuraminidase from *Clostridium perfringens* type VI (Sigma-Aldrich, USA) diluted in PBS to a final concentration of 0.2 U/ml, for 1 h at 37°C. Peroxidase-conjugated secondary antibodies (goat polyclonal anti-mouse-HRP 1:2000 for CDX2, TetR and MUC2 and goat polyclonal anti-rabbit-HRP 1:2000 for actin (Santa Cruz Biotechnology, USA) were used and developed with the ECL detection kit (Bio-Rad Laboratories, USA).

RNA extraction

Total RNA was extracted using RNeasy kit (Qiagen, GE) as recommended by the manufacturer.

Transcriptomics (RNA-Seq)

Polyadenylated RNA was isolated from total RNA using standard protocols (Dynabeads mRNA Direct Micro Kit, Ambion/Life Technologies, USA) starting with 5 μg total RNA. Library preparation for NGS using Ion Torrent technology (Life Technologies, USA) was carried out according

to the manufacturers recommendations (Ion Total RNA-Seq Kit v2, Ion torrent/Life Technologies, USA), and sequenced using an Ion Proton system (Life Technologies, USA). Quality control, quantification of RNA and libraries was carried out using using Agilent RNA 6000 Nano Kit or Agilent High Sensitivity DNA kit and Agilent Bioanalyzer (Agilent Technologies, USA). Sequencing reads were mapped to hg19 and bioinformatics analysis was conducted using CLCs Genomic Workbench (CLC bio/Qiagen, DK), STARBowtie2 followed by Cufflinks for expression analyses. In brief, ∼50 ng polyA RNA was fragmented down to 100–300 base fragments using RNaseIII for 1–3 min followed by adapter ligation, amplification for 9–14 cycles and barcoding using Ion Express RNA-Seq Barcode kit (Ion Torrent/Life Technologies, USA). The final library fragment size and concentration was determined by Agilent Bioanalyzer analysis followed by template preparation using Ion PI Template OT2 200 Kit v3 (Life Technologies, USA) and Ion One Touch System followed by NGS on an Ion Proton system using Ion PI™ Chip Kit v2. In general, two transcriptome libraries were barcoded and analyzed on one Ion PI v2 chip. Sequencing depth for the RNA-Seq data sets ranged from 34.5 to 44.6 mill reads/sample with 95.4–97.7% mappable reads to the hg19 reference data base. Mean read length for transcriptome RNA-Seq data ranged from 116 to 136 bp.

Genome-wide analysis of CDX2-binding sites by ChIP-seq.

LS174T cells grown for five days in a 30 × 30 cm culture dish were cross-linking and sonicated as described previously to generate fragments of ∼0.2 to 1.2 kb. Hereafter, the ChIP-protocol was performed as previously reported (34). Briefly, immunoprecipitation was done in four replicates and performed overnight at 4°C using specific antibodies to human CDX2 (CDX2-88, Biogenex, USA) and the influenza hemagglutinin (HA) epitope (rabbit polyclonal α-HA; Santa Cruz Biotechnology Inc, Heidelberg, Germany), used as a negative control. Immunocomplexes were recovered with 50 μl protein A/G beads (Invitrogen/ThermoFisher Scientific, USA). Verification of the enrichment in the CDX2-immunoprecipitated DNA samples was done by qPCR analyzing the DNA level from known CDX2-targets in the CDX2-immunoprecipitated samples and comparison it to level in the negative control (HA-immunoprecipitated DNA).

In order to increase the depth of the analysis, the ChIP library established was deep sequenced using Ion Proton instrumentation generating an approximate ≈54 million reads depth. Library preparation for NGS was carried out with the Ion Xpress Plus gDNA Fragment Library Kit according to User Bulletin 4473623: Ion Chipseq Library Preparation on the Ion Proton System (Life Technologies, USA), starting with 10 ng DNA. Quality control of the libraries was carried out using using Experion DNA 1K analysis kit and the Experion System (Bio-Rad Laboratories, USA). In brief, 10 ng DNA was end-repaired and ligated to barcode adaptors from the Ion Xpress Barcode Adapters 1–16 Kit (Life Technologies, USA), followed by nick repair and amplification for 18 cycles. The amplified libraries were subjected to two rounds of bead capture with the Agencourt AMPure XP Kit (Beckman Coulter, USA) to size-select fragments ∼160–340 bp in length. The final library concentration was determined by Qubit analysis (ThermoFisher Scientific, USA). The barcoded ChIP- and IP control-DNA-libraries were handled similarly to the RNA-Seq library and analyzed on a single Ion PI™ Chip v2. For the ChIP- and IP control-DNA-libraries a total of 54.5 mill reads (21.8 and 32.7 mill respectively) were obtained and 98.5% were mappable to the reference hg19 data base with a mean read length of 130 bp. The detection of CDX2 ChIP seq peaks and the location of the closest gene was performed using CisGenome version 2 (35).

Immunodetection by fluorescence-activated cell-scanning (FACS) analysis

Cells were trypsinized at different time-points after induction with Dox or after Dox removal and washed twice with PBS. After being centrifuged at 1200 rpm for 7 min, CDX2 KI cells were fixed in 4% PFA for 20 min., washed in PBS and permeabilized in cold methanol for 15 min. After washing with PBS, samples were incubated with the primary antibody (CDX2-88 clone, BioGenex, 1:500.) for 1 h at 37°C. Cells were then stained with FITC-conjugated AfinityPure immunoglobulin antimouse IgG (Jackson ImmunoResearch Laboratories, USA) diluted 1:100 in 0.05% BSA in PBS and then subjected to FACS in a FACS ARIA III (BD BioSciences, USA). T2EYFP KI cells were immediately scanned by FACS after cell trypsinization.

Immunocytochemistry

Cell slides were fixed in cold acetone for 15 min. For TetR anti-TetR cells were incubated overnight at 4°C with the primary antibodies (9G9 clone, Clontech/Takara, USA). Negative controls were performed by omission of primary antibodies. After washing, a rabbit anti-mouse Ig FITC-labeled secondary antibody (Dako A/S, DK) diluted 1:100 in PBS with 5% BSA was added for 45 min, protected from light. DAPI was used as a nuclear counterstain and slides were mounted in Vectashield mounting media (Vector Labs, USA).

For EYFP visualization on the induced T2EYFP KI clones, trypsinized cells were fixed in cold acetone for 10 min, DAPI was added and slides mounted. Samples were examined under a Zeiss fluorescence microscope equipped with DAPI and FITC interference filters. Images were acquired using a Zeiss Axioskop 2 and an AxioCam MR3 camera and Zeiss Application Suit software.

Real time *in vivo* video material

For the time lapse video, 1×10^6 CDX2$^{-/-}$2XindT2EYFP cells were seeded in a well of a six-well plate and 0.5 μg/ml of Dox were added after 24 h. Time lapse images started to be acquired under dark field immediately after Dox addition and every 20 min during 24 h using a Leica DMI 6000 timelapse microscope equipped with FITC interference filter and with CO_2 supply.

RESULTS

Precise integration of inducible transcriptional elements (PrIITE)

CDX2 knock-out by targeted integration of Tet3G transactivator elements. In order to eliminate the adverse cellular effects seen from prolonged exposure to Dox (25,26), we designed a strategy combining stable and controlled TeT-On transactivator/TeT3G expression with disruption of a target gene. First we established a Tet3G expressing LS174T cell line, in which the TET3G elements were targeted to *CDX2* exon1 CDS, thereby abrogating the endogenous gene function. Targeting *CDX2*, a master differentiation transcription factor (36), allowed us to determine the efficacy of the system in an unprecedented CDX2 knock out cell system (Figure 1). Multiple attempts at homologous recombination driven donor integration at the CDX2 exon 1 locus based on CDX2 ZFNs and a TET3G donor template flanked by 700bp homology arms did not give rise to any correctly targeted clones, likely due to low homologous repair capacity of LS174T cells used in this study (data not shown). *pCDX2-pCMV-TET3G-ObLiGaRe* donor vector (Figure 1A and Supplementary Figure S1A) was co-transfected together with *CDX2* CompoZr ZFN plasmids into LS174T cells, whereby *pCMV-TET3G transactivator* donor vector was specifically integrated into exon1 of *CDX2*, as illustrated in Figure 1A and Supplementary Figure S1A and B, by ObLiGaRe driven recombination (see Material and Methods section for details). Three clones (5E2, 3D6 and 5B5) were shown to include correct integration of the *p*CMV-TET3G construct and constitutively express the Tet3G transactivator (Supplementary Figure S1C), one clone, 5B5, was mono allelically targeted (designated $LS^{CDX2+/-}$) and CDX2 protein expression maintained at levels similar to the parental LS174T wt cell line and in the bi-allelic targeted *CDX2* knock-out (KO) clones, 3D6 ($LS^{CDX2-/-}$#1) and, 5E2 ($LS^{CDX2-/-}$#2), *CDX2* expression was completely abrogated (Supplementary Figure S1C and D). Constitutive Tet3G expression and abrogation of CDX2 was maintained after prolonged cell passaging and after freeze/thawing lack of CDX2 was correlated with significantly decreased protein levels of a well-know CDX2 target gene, *MUC2* (37) (Supplementary Figure S1C and D). Lastly, we tested the functionality of the mono and bi–allelically targeted *CDX2* KO cells by cellular transfection with known CDX2 enhancer reporter constructs and could show dose dependent decreased expression of a *HNF4A* reporter and >10-fold reduced expression of a *HEPH1* reporter (38) in the *CDX2* KO cells (Supplementary Figure S3).

We thus show, that *CDX2* KO by ZFN-mediated site specific *p*CMV-TET3G integration in LS174T cells was effectively accomplished and that stable Tet3G expression was maintained over time.

Safe harbor targeted integration of inducible T2EYFP elements. Having shown that both $LS^{CDX2-/-}$#1 and #2 display similar transactivator expression levels, an inducible T2EYFP model system was established in $LS^{CDX2-/-}$#1. The inducible *p*TRE3G-T2EYFP-*ObLiGaRe* donor vector was targeted to the safe harbor *AAVS1* locus by co-transfection with *AAVS1* CompoZr ZFN's into the $LS^{CDX2-/-}$#1 cells as illustrated in Figure 2A. Multiple correctly targeted indT2EYFP clones were obtained and two of these selected for further detailed analysis, one mono-allelic T2EYFP targeted clone 7E9 (hence forward referred to as $CDX2^{-/-}$1XindT2EYFP) and another bi-allelic targeted clone 5G8 (henceforward referred to as $CDX2^{-/-}$2XindT2EYFP), Supplementary Figure S2. Both clones maintained constitutive Tet3G expression and absence of CDX2 (Figure 2B).

LS174T indT2EYFP PrIITE cells display no leakiness and Dox induction is fast and reversible. We first aimed to determine the leakiness of the inducible T2EYFP reporter in the $CDX2^{-/-}$1xindT2EYFP and $CDX2^{-/-}$2xindT2EYFP clones by FACS and immunofluorescence, which showed undetectable fluorescence in uninduced PrIITE cells (Figure 2C, D and E). Upon induction, immunofluorescence displayed the expected peri-nuclear reactivity, consistent with the correct sub-cellular localization of T2EYFP fusion protein in the Golgi apparatus (39) (Figure 2E).

We next determined the induction dynamics over time by *in vivo* live imaging, demonstrating that the induction dynamics is very rapid. T2EYFP expression is detectable within 4 h post induction and all cells display full expression of the reporter after 8–10 h (Supplementary video material). Importantly, the induction dynamics is fully synchronized and all cells simultaneously become positive for the T2EYFP reporter. To elucidate how Dox dosage might influence on the dynamics of the system, induction was performed using variable Dox concentrations ranging from 0 to 5 μg/ml. As shown in Figure 2C, modest induction was observed with dox concentrations <5 ng/ml, whereas significant induction required Dox concentration >10 ng/ml. Interestingly, these results suggest, that the induction potential is related to the number of inducible elements integrated into the system. For instance, using a 10 ng/ml Dox dosage rendered >90% of the $CDX2^{-/-}$2XindT2EYFP cells EYFP reporter positive, whereas <50% $CDX2^{-/-}$1xindT2EYFP cells were found positive for this dosage.

Next, the reversibility of the PrIITE cell system was quantified by FACS analysis of the EYFP expressing cell population after Dox removal from the media (Figure 2D). As observed, induction was fully reversible and after 10 days post Dox removal, all cells in the population returned to non-expressing levels. However, the dynamics in reversibility was found to be dependent on both number of inducible elements integrated and the Dox dosage used. Clearly reversal to non-expressing levels was achieved rapidly and within 5 days for $CDX2^{-/-}$1XindT2EYFP and slower for $CDX2^{-/-}$2xindT2EYFP using the lower 50 ng/ml induction dosage.

We thus demonstrate that the PrIITE cell system displays no leakiness, is fully reversible, and that the induction dynamics occurs fast and in a synchronized manner and require below cellular stress inducing Dox concentrations.

LS174T indCDX2 PrIITE cells show no leakiness and ectopic CDX2 expression is tightly controlled requiring minimal Dox induction levels. The role of CDX2 in intesti-

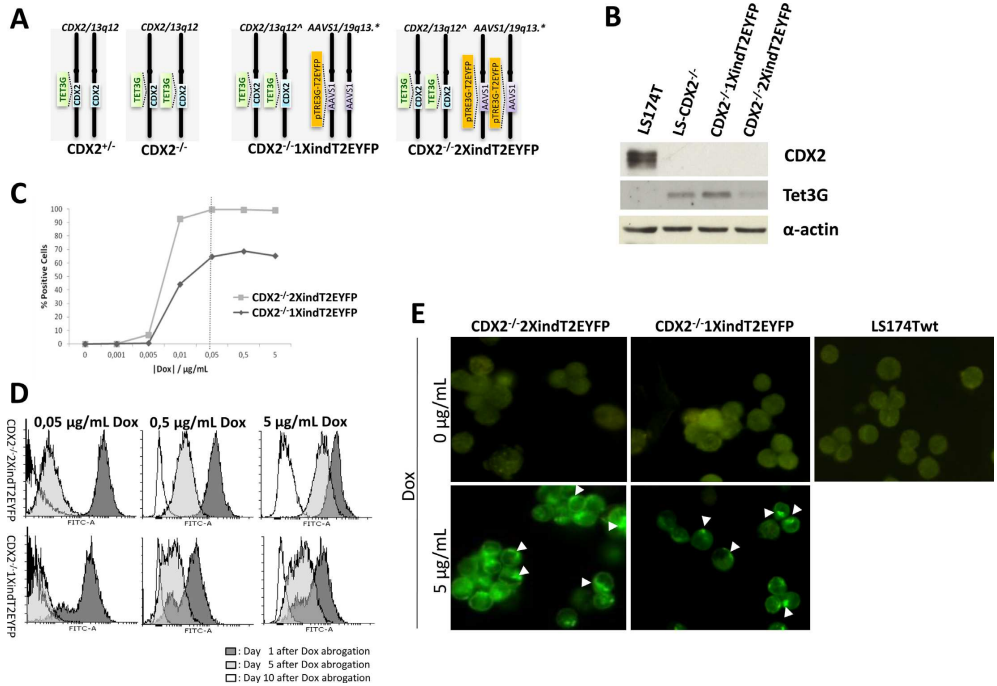


Figure 2. Overview of the proof of principle system using inducible T2EYFP reporter in isogenic LS174T-PrIITE cells. (**A**) First pCMV-Tet3G was targeted to *CDX2* exon1 of LS174T cells generating mono- or bi-allelic CDX2 KO cells, $CDX2^{+/-}$ and $CDX2^{-/-}$ respectively. Next, cells were re-targeted with *pTRE3G-T2EYFP ObLiGaRe-donor vector* (pTRE3G-T2EYFP) directed to the safe harbor *AAVS1* locus generating $CDX2^{-/-}$1XindT2EYFP and $CDX2^{-/-}$2XindT2EYFP cells. (**B**) Western blot analysis of LS174wt, $CDX2^{-/-}$, $CDX2^{-/-}$1XindT2EYFP and $CDX2^{-/-}$2XindT2EYFP cells show absence of CDX2 and stable TeT3G expression. (**C**) FACS based determination of percentage indT2EYFP positive cells after induction with different Dox concentrations. The graph shows that induction is dose and copy number dependent in $CDX2^{-/-}$1XindT2EYFP and $CDX2^{-/-}$2XindT2EYFP PrIITE cells. Approximate 100% T2EYFP positive cells was achieved using >0.005 µg/ml Dox concentration for $CDX2^{-/-}$2XindT2EYFP. Maximum 70% positivity was achieved regardless of the higher Dox induction concentrations used for $CDX2^{-/-}$1XindT2EYFP cells. Importantly, $CDX2^{-/-}$2XindT2EYFP cells show complete absence in T2EYFP fluorescence in uninduced cells thus, no leakiness of T2EYFP LS-PrIITE cells was detectable. Critical dox concentration above which adverse cellular stress (24–26) is caused is indicated by a dotted gray line. (**D**) Reversibility of induction as determined by FACS analysis of fixed cells 1, 5 or 10 days post 48 h induction with variable Dox induction concentrations. Notably, reversibility was copy number and dox dependent and full reversibility within 10 days was only achieved for $CDX2^{-/-}$1XindT2EYFP PrIITE cells using the lower 0.05 µg/ml Dox concentration. (**E**) Sub cellular localization of inducible T2EYFP as determined by fluorescence of fixed trypsinized indCDX2 PrIITE cells with or without dox induction. Arrow heads indicate expected Golgi localization of T2EYFP.

nal gene regulation is well known. However, to the best of our knowledge no studies so far described the effect of complete ablation of *CDX2* from the cellular genome. In a single study *CDX2* has been suggested as a lineage survival oncogene when amplified in colorectal cancer (40). In order to distinguish endogenous from ectopically expressed *CDX2* transcripts the inducible *CDX2* open reading frame was codon optimized (hereafter referred to as indCDX2). *pTRE3G-indCDX2-ObLiGaRe* donor vector was targeted to the *AAVS1* safe harbor site of $LS^{CDX2-/-}$#1 cells generating clones 7D9 ($CDX2^{-/-}$1XindCDX2) and 6D6 ($CDX2^{-/-}$2XindCDX2) possessing one or two inducible codon optimized CDX2 copies respectively (Figure 1B and Supplementary Figure S4). Constitutive Tet3G expression was maintained in targeted cells and $CDX2^{-/-}$2XindCDX2 or $CDX2^{-/-}$1XindCDX2 cell induction dynamics over 24 h were similar (Figure 3A and B). To test the reversibility of the CDX2 PrIITE system built, we induced with a range of Dox concentrations (0–4 µg/ml) for 48 h followed by removal of Dox from the medium. As expected, no expression of indCDX2 was revealed in non-induced cells followed by fast dose dependent Dox induction dynamics (Figure 3C). Importantly, for the lowest concentration of Dox used (0.004 µg/ml), the induction dynamics for the $CDX2^{-/-}$ 1XindCDX2 clone was slower than for $CDX2^{-/-}$2XindCDX2 (Figure 3C). Of notice, the indCDX2 protein levels, reached for $CDX2^{-/-}$1XindCDX2 with 4 ng/ml Dox dose, were similar to the endogenous CDX2 levels detected in LS174Twt cells (Supplementary Figure S4E) while for $CDX2^{-/-}$2XindCDX2 induction reached 120× higher levels relative to LS174Twt levels (Supplementary Figure S4E). In contrast to the dynamics in reversibility for indT2EYFP KI clones (Figure 2), reversibility of indCDX2 induction appeared faster and required very modest Dox induction dosages in the sub cell stress inducing nanogram/ml range.

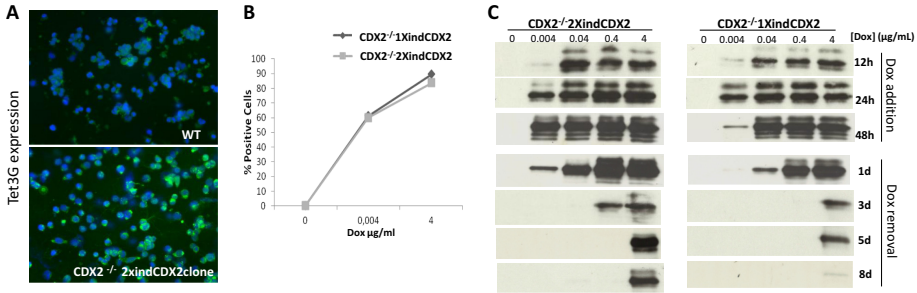


Figure 3. Induction dynamics of isogenic inducible CDX2 LS174T PrIITE cells. (**A**) Anti-TeT3G transactivator immunofluorescence in CDX2$^{-/-}$2XindCDX2 cells and not in LS174T wt cells. (**B**) 24 h induction kinetics of indCDX2 using different Dox concentrations. >60% of CDX2$^{-/-}$1XindCDX2 and CDX2$^{-/-}$2Xind CDX2 cells become indCDX2 positive after 24 h 0.004 μg/ml Dox induction. Use of higher Dox concentrations higher percentage of cells become CDX2 positive after 24 h Dox treatment. (**C**) Dynamics in reversibility of the PrIITE system using different Dox induction concentrations. CDX2$^{-/-}$2Xind CDX2 (left panel) and CDX2$^{-/-}$1Xind CDX2 (right panel) cells were induced 48 h with the Dox concentrations indicated above the panels. Hereafter Dox was removed and post induction indCDX2 protein levels were determined by Western blot analysis 1, 3, 5 or 8 days post Dox removal. Maximum indCDX2 induction levels were achieved using 0.004 μg/ml Dox concentration for CDX2$^{-/-}$2XindCDX2 whereas 0.040 μg/ml Dox was required for CDX2$^{-/-}$1Xind CDX2.

Based on western blot analysis indCDX2 protein was undetectable within 3 days post Dox removal for both mono- and bi-allelic clones (Figure 3C).

Validating the tightness in inducibility of the PrIITE by deep transcriptome analysis. To confirm the tightness of the PrIITE system and in order to determine the CDX2 downstream target genes, LS174Twt, uninduced and induced CDX2$^{-/-}$2XindCDX2 or CDX2$^{-/-}$1XindCDX2 cells were RNA-seq transcriptome profiled (Figure 4A, and Supplementary Figure S5). To ensure that the isogenic PrIITE cells generated were representative of the mosaic moderately well differentiated LS174Twt cell population (41), all comparative analysis here and in the following sections included LS174Twt cells. Approximately 10 000 genes with RPKM >1 were shown to be expressed in the wt cells and the expression profiles for the mucin genes *MUC2* and *MUC5AC* correspond to the profiles previously reported for LS174Twt cells with high expression for the former and low expression for the latter (42,43) (Figure 5B). CDX2 target genes were defined by the following criteria; genes with RPKM >1 in wt cells where RPKM values in uninduced PrIITE cells (*CDX2*$^{-/-}$) were reduced >4× (4×↓) and increased >2× (2×↑) after re-induction relatively to the uninduced RPKM values. In depth analysis of the RNASeq reads that mapped to the human exon1 *CDX2* gene locus, including the region across the *CDX2* ZFNs cutting site, confirmed the presence of out of frame transcripts possessing deletions in CDX2$^{-/-}$2XindCDX2 exon1, while the heterozygous CDX2$^{-/-}$1XindCDX2 clone expressed intact endogenous *CDX2* transcripts present at ≈40% of the LS174Twt transcript levels (Supplementary Figure S5A). Analysis of the unmapped RNA-seq reads confirmed the presence of the Tet3G transactivator transcripts in CDX2$^{-/-}$ cells (Supplementary Figure S5B) and importantly, no indCDX2 transcript was detected in the uninduced state of the CDX2$^{-/-}$2XindCDX2 clone (Supplementary Figure S5B), thus confirming the biochemical and immunofluorescence results obtained previously.

Taken together, by biochemical analysis, immunofluorescence and deep transcriptome analysis, we have shown that the PrIITE system allows for absolute control of inducible and reversible gene expression of CDX2 and it's downstream target genes.

Validating the transcriptome identified CDX2 target genes by genome-wide ChIP-seq analysis. Reasoning that the RNA-seq identified CDX2 target genes could be direct or indirect targets of CDX2, we wanted to validate the RNA-seq data set by genome wide chromatin immunoprecipitation sequence analysis (ChIP-seq). We thus used ChIP-seq to identify true CDX2 binding elements in the vicinity of the RNA-seq identified CDX2 target genes. ChIP was performed on LS174Twt cells using a well-known anti-CDX2 monoclonal antibody essentially using a previously reported procedure (38). A LS174T non-immunoprecipitated library was deep sequenced in parallel and used as ChIP-seq background control. An additional requirement was added to the RNA-seq criteria for positive scoring of primary CDX2 target genes, in that a ChIP-seq peak (CDX2 binding element) were to be present within or 1 kb up or downstream of the target gene. The combined RNA- and ChIP-seq criterion identified 31 direct CDX2 target genes (Figure 4B and Supplementary Table SI). Among these genes several known CDX2 regulated targets were found, such as *TFF3*, *MUC2* and *CDX2* itself, but several novel targets were also found, including *MUC5B, MUC5AC*, *MUC6*, *GPA33* and *LDLR* (Figure 5). Furthermore, *in vitro* promotor analysis confirmed the CDX2 dependent regulation of the enhancer identified in the novel CDX2 target gene *GPA33* (44) (Figure 6) and thus, substantiated both our RNA- and ChIP-seq findings. Interestingly, a significant intergenic CDX2 binding element was identified in the 300 Mb 11p15 region containing the mucin genes *MUC6*, *MUC2*, *MUC5B* and *MUC5AC* (Supplementary Figure S6A). Clearly our transcriptome results reveal a significant concerted down-regulation of this gene cluster and suggest CDX2 as a 'locus control gene' (Figure 5B). As expected, the non CDX2

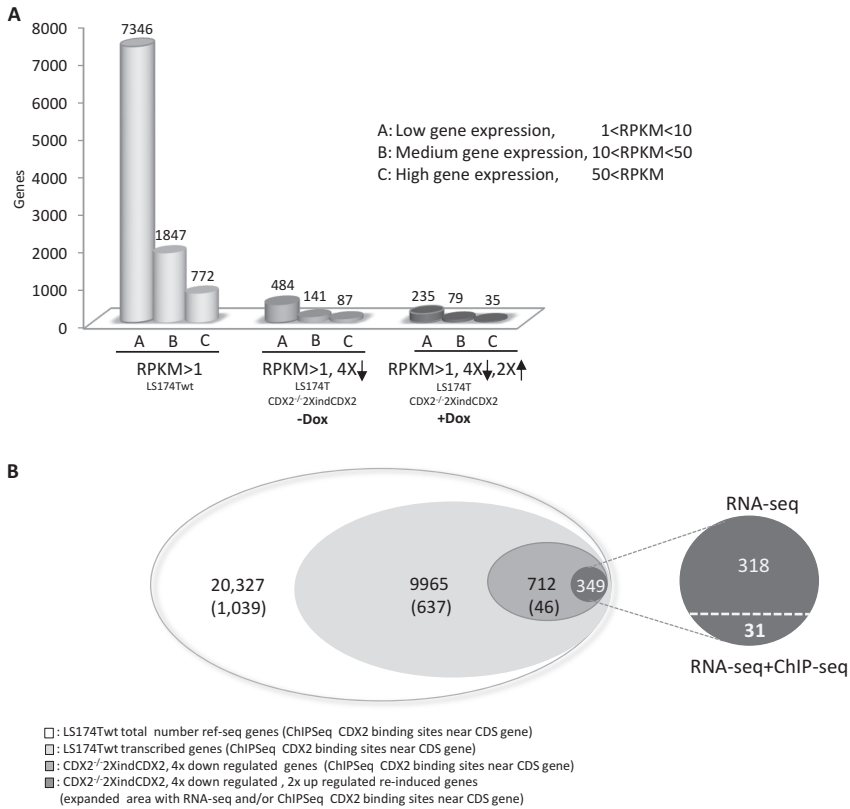


Figure 4. Deep RNA-seq and genome wide ChIP-seq analysis of LS174Twt and induced or uninduced CDX2$^{-/-}$2XindCDX2 PrIITE cells. (**A**) Of the 20 327 human genes included in the analysis around 9965 were found to be expressed with RPKM >1, the majority hereof (7346) were low expressed genes with 1 < RPKM < 10 (A), followed by medium expressed genes (1847) with 10 < RPKM < 50 (**B**) and highly expressed genes (772) with 50 < RPKM (**C**). In uninduced CDX2$^{-/-}$2XindCDX2 (−Dox) PrIITE cells (lacking CDX2), 712 genes expressed in LS174Twt were 4-fold down regulated and after indCDX2 re-induction (+Dox) 349 of these genes were 2-fold upregulated and thus suggested to be CDX2 downstream target genes. (**B**) In order to determine to what degree the genes identified represent direct CDX2 target genes, we included a LS174T ChIP-seq data set and seeked for the presence of potential CDX2 regulatory elements in the vicinity of the RNA-seq suggested CDX2 target genes. Inclusion of this criteria (numbers shown in parenthesis) in addition to the RNA-seq criterion defined above resulted in a dramatic decrease in potential direct CDX2 targets going from 349 to 31 genes. The expanded view at right depicts the number of targets identified using the RNA-seq or RNA-seq+ChIP-seq criteria. In the latter case the targets indicated below the stippled line are listed in Supplementary Table SI.

regulated mucin genes *MUC1* and *MUC7* did not possess any CDX2 binding elements and did not respond to CDX2 changes (Figure 5B and Supplementary Figure S6B). Previously, we have identified CDX2 as a regulator of glycosyltransferase encoded *ST6GALNAC1* expression (45). However, in this study, no CDX2 binding elements were identified in the *ST6GALNAC1* locus (Supplementary Figure S6B), which suggests that CDX2 regulates the expression of this gene by an indirect effect.

The results also show that several known CDX2 targets such as *FUT2* (46), *B3GALT5* (47), *ALP1* (48) and *SI* (28) in LSCDX2 PrIITE cells were unaffected by CDX2 changes (Figure 5B and Supplementary Figure S7) and notably, in the latter three cases these genes were shown not to be expressed in LS174Twt cells. Of notice, the only glycogenes (49) otherwise affected by CDX2 were *FUT3* and *GALNT1* which in the latter case responded inversely to CDX2 changes (Figure 5B).

In order to assess to what degree the global reversibility in gene expression for LSCDX2 PrIITE cells was retained, RPKM values from LS174Twt and LSCDX2 PrIITE cells exposed to different Dox concentrations ranging from 0 to 40 ng/ml were plotted and the regression line for each data calculated (Figure 7 and Supplementary Figure S8). The results show that global reversibility of CDX2 controlled gene expression was profound, and in many cases full reversibility was observed when the lowest 4 ng/ml Dox concentration was used with CDX2$^{-/-}$2XindCDX2 cells. Clearly, the combined RNA-seq and ChIPseq filtering strategy employed improved the specificity in defining true CDX2 regulated genes in our data set (Figure 7C). We speculate, that the observed highest normalization levels of

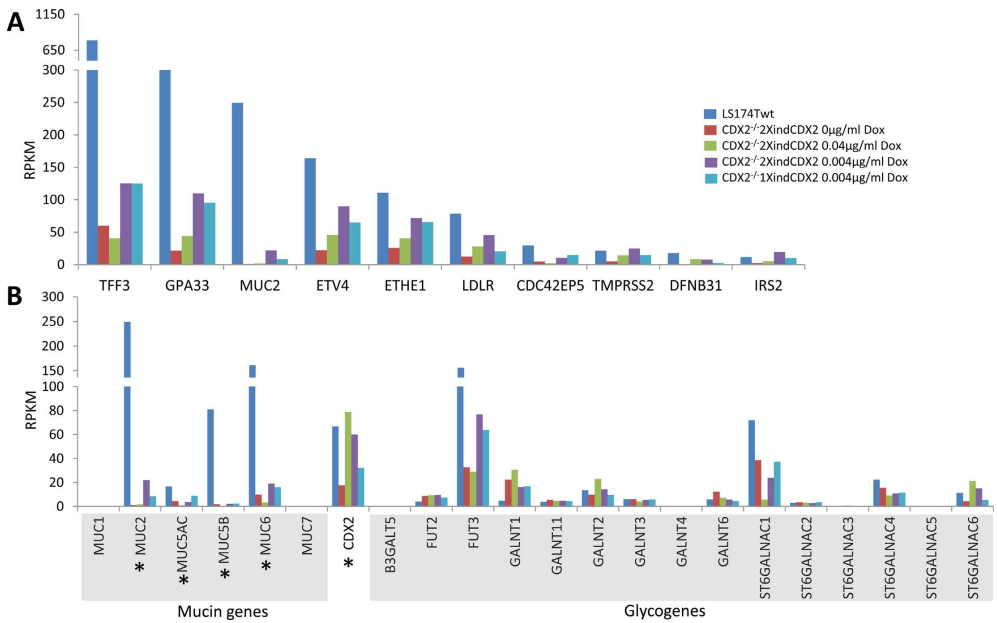


Figure 5. CDX2 regulated target genes. (**A**) Based on the included RNA-seq and ChIPseq criteria (see Figure 4 and Supplementary Table SI) top 10 primary targets downstream of CDX2 with most abundant transcripts (highest RPKM values) were ranked. Well known CDX2 target genes were identified such as *TFF3* and *MUC2*, but novel targets were also identified including *GPA33* and *LDLR*. (**B**) Careful analysis of 208 glycogenes (49) controlling the cellular glycophenotype and 23 mucin genes revealed that *MUC5B*, *MUC6* and *MUC5AC* were controlled by CDX2, whereas *MUC1* and *MUC7* did show no correlation with CDX2 expression. No CDX2 correlation with *FUT2* and *B3GALT5* could be demonstrated, which is in contrast with previous reports. On the other hand, *ST6GALNACI* and *-VI* were shown to correlate with CDX2 expression although no CDX2 binding element could be identified within the vicinity of the CDS of these genes. Interestingly, an inverse correlation was observed for *GALNT1* regulation similar to *SOX2* (see Supplementary Figure S7), with upregulation in the absence of CDX2 expression. Interestingly, the novel CDX2 target *GPA33* gene product has been shown to be post translationally modified by the *GALNT1* encoded enzyme. Genes with ChIP-seq identified CDX2 binding elements in the gene locus vicinity are indicated by an *.

re-induced CDX2 for 4 ng/ml Dox may be due to the role of CDX2 as a molecular rheostat potentially controlling: (i) self-renewal as indicated in Figure 5 and/or (ii) regulated expression of downstream target genes within discrete and narrow CDX2 expression limits (50). Lastly, in order to determine if the genetic manipulations had changed the differentiation state and stemness of CDX2 PrIITE cells generated, we determined the expression profile of 24 marker genes before and after genetic manipulation and/or CDX2 re-induction (Supplementary Figure S7). In all cases (except for *MUC6*), the gene expression profiles of genetically manipulated cells relative to LS174Twt was retained and displayed a distinct intestinal expression pattern. Interestingly, the stemness related gene, *SOX2*, is not expressed in the wt cells, but strongly expressed upon CDX2 KO and clearly decreased upon reinduction. This may suggest that also stemness regulation is CDX2 dependent. The CDX2 targets identified in our data-set is to some extent in agreement with previous findings in a related cellular background (38) (Supplementary Table SI). Of note, the karyotype of the LS174T cells used in our study was determined and found to be in complete agreement with the originally near normal published karyotype (41) (Supplementary Figure S9).

DISCUSSION

Gene regulatory mechanisms are complex and profit from the development of systems in which gene and protein expression is tightly controlled. Tet-regulated expression systems have been widely used for inducible protein expression in mammalian cells. However, current optimized Tet-on systems are still hampered by residual levels of gene expression in the uninduced state in both transiently and stably transfected cells (18,21,22). We hereby demonstrate that the PrIITE system resolves all issues related to leakiness and importantly requires use of tetracycline dosages below cell stress inducing concentrations (24). We first prove the PrIITE concept using a Golgi targeted EYFP model reporter (31). With this tightly controlled inducible reporter system at hand, we were able to demonstrate, to the best of our knowledge for the first time, that leakiness relates to increased expression of the TET3G transactivator element and not increased presence of the inducible GOI, in this case the T2EYFP reporter, Supplementary Figure S10. We furthermore show that stable random integrated pCMV-TET3G transactivator elements causes leaky expression in a substantial sub-population of the cells and we speculate that this sub-population of cells enable uncontrolled Dox induction levels beyond what is achievable in the tightly

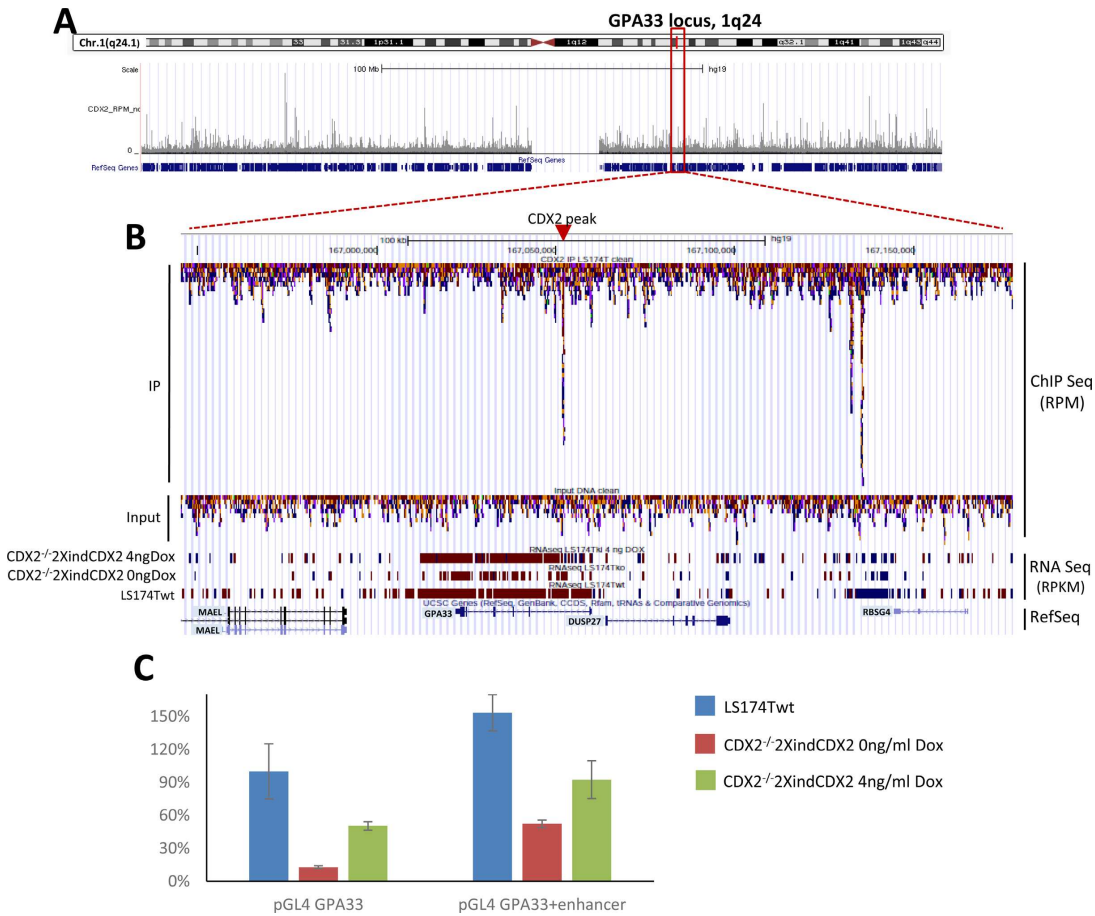


Figure 6. Identification and validation of the novel CDX2 target gene, *GPA33*. (**A**) Panel shows chromosome 1 coverage of CDX2 ChIPseq reads (CDX2IP) and control (input), calculated in 300 bp sliding windows as RPM (reads per million). For each window, input RPM was subtracted from CDX2IP RPM and normalized values were saved as a bedgraph file for visualization in the UCSC Genome Browser. Major CDX2 binding sites (peaks) are shown with indication of the *GPA33* locus framed in red box. (**B**) Expanded display of the GPA33 locus, showing detailed identification of CDX2 ChIPSeq peaks (ChIPseq/RPM) in both LS174T CDX2IP and control input samples. Major *GPA33* CDX2 binding element is indicated by red triangle (CDX2 peak). RNA-seq panel displays the density of individual reads in CDX2$^{-/-}$2XindCDX2 with or without Dox induction (4 or 0 ng respectively) and LS174Twt cells. RefSeq gene panel displays the genomic organization of neighboring genes in the *GPA33* locus. (**C**) *In vitro* promoter analysis of the 0.6 kb CDX2 enhancer element identified in the *GPA33* intron. pGL4-GPA33 and pGL4-*GPA33*enhancer represent the reporter plasmid without or with the enhancer element included. Reporter expression is clearly dependent on both GPA33 CDX2 enhancer and CDX2 presence.

controlled PrIITE cells generated, Supplementary Figure S10B. We next exemplify the potential of the system by clarifying the role of CDX2 in intestinal cells as a consequence of controlled *CDX2* expression. This model system shows that CDX2 predominantly acts as a transcriptional activator and only a limited number of genes seemed to be repressed by CDX2 presence, SOX2 being an example hereof (Supplementary Table SII). We had anticipated that CDX2 targeting in LS174T cells could become problematic due to its proposed essential cellular function and pre-designed our targeting strategy for dealing with essential genes. This was accomplished by incorporating a 'landing pad' in the EPB71 AAVS1 donor integration vector, just downstream of the inducible gene of interest (Supplementary Figure S11). The 'landing pad' encodes the Safe Harbor #1 (SH1) sequence derived from the CHO genome, that has successfully been utilized by us (51) and others (52) for ZFN mediated target integration in CHO cells and thus represents a unique site when integrated in human cells devoid of this sequence. This allows for subsequent donor target integration into the SH1 site. Establishment of a PrIITE cell model for an essential gene can thus be accomplished, by targeting of both the TET3G transactivator and inducible pTRE3G codon optimized GOI transcriptional elements to

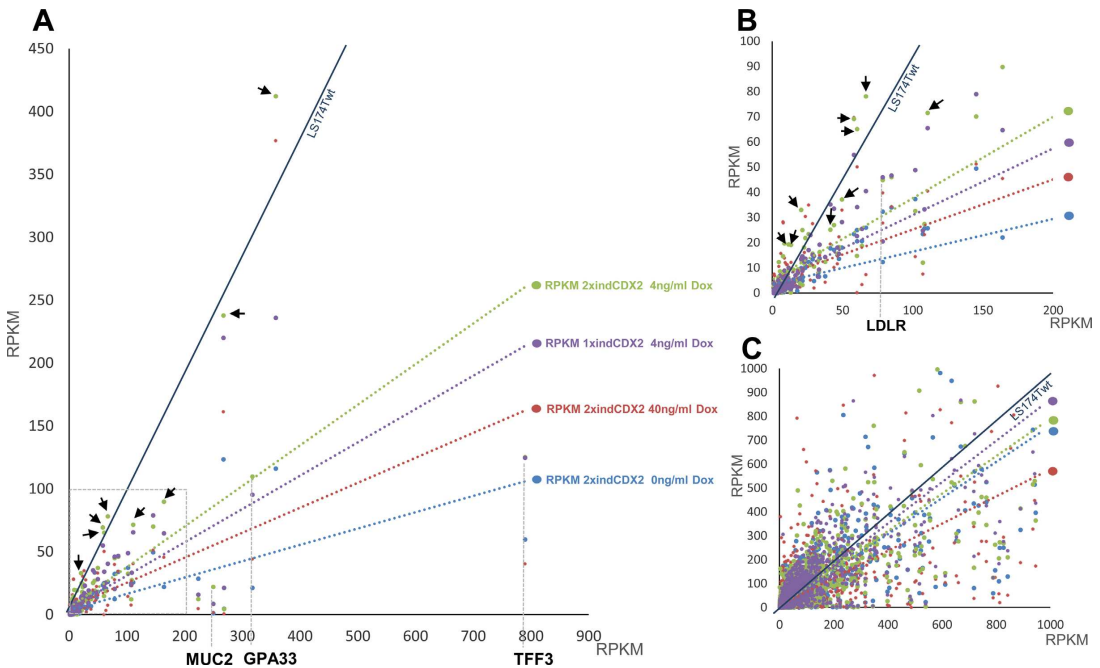


Figure 7. Global Dox induction reversibility of the PrIITE cells relative to LS174Twt cells. In order to include enough data points in the analysis, filtering criteria included all genes with RPKM >1, that were 2× reduced in CDX2$^{-/-}$ KO cells and had a CDX2 binding element/peak in the gene locus. RNA-seq data (RPKM) from wildtype LS174T was plotted against RPKM values from uninduced CDX2$^{-/-}$2XindCDX2 cells without CDX2 expression (blue dotted line, 2XindCDX 0 ng DOX) and under different conditions of CDX2/Dox induction (ng/ml) (green dotted line, 2XindCDX2 4 ng DOX; purple dotted line, 1XindCDX2 4 ng DOX, red dotted line, 2XindCDX2 40 ng DOX). The regression line for each data set is plotted. The dark blue solid line represents the expression levels in LS174Twt cells. (**A**) Global display of RNA-seq RPKM values for the samples analysed. RPKMs for the known CDX2 target genes *MUC2 and TFF3* and the novel target gene, *GPA33*, are indicated by grey dotted vertical lines. Box with gray dotted line is area shown in panel B. Representative examples of genes displaying full reversibility are indicated with arrow heads. (**B**) Zoom from dotted box from panel A with indication of RPKM values for the novel CDX2 target gene, *LDLR*. Representative genes for which complete reversibility was observed are indicated by arrow heads. (**C**) Global display of RNA-seq RPKMs for genes fulfilling the RPKM >1 and 2× reduction filtering criteria but without fulfilling the criteria of a CDX2 binding element/peak in the gene locus. Note the unordered dispersed display of RPKMs relative to the LS174Twt regression line, suggesting that a large proportion of the CDX2 target genes only based on RNA-seq data are not direct targets of CDX2.

the AAVS1 locus, ex. by AAVS1 targeting of the former gene followed by SH1 targeting of the later gene. Hereafter, precise target inactivation of the endogenous gene of interest in the presence of Dox would ensure continued expression of the inducible codon optimized version of the essential GOI. In this project, we did not take advantage of this option, since in the LS174T cellular model CDX2 has not proven to act as a lineage-survival oncogene in contrast to what is observed in other cell line models (40).

Our results also show that LS174T cells have a poor capacity to integrate homology arm flanked donor templates, which is likely attributed to the lacking HR capacity of these and other cell lines used (53–55). In general, precise donor integration remains a challenge in the field and in spite of the fact that CRISPR efficiently induces double stranded breaks, CRISPR mediated donor integration importantly comes at the expense of: (i) 3–20-fold higher indel formation rates at the non-targeted allele (56,57) and (ii) significant off-target/random integration frequencies (58). These concerns were evident in a recent *in vivo* study based on CRISPR target integration of inducible TRE3G promoter elements, where 'leaky' expression in 33% of clones analyzed was identified (59). In light of these CRISPR related issues, the ObLiGaRe targeting strategy by Marcello Maresca *et al.* remains to be an efficient and reliable target integration strategy for most cells (60–63), including HR compromised cells such as CHO cells (64,65) and colorectal cancer lines as described in this study and by Marcello Maresca et al.

A whole set of novel genes, including *GPA33* (44), *LDLR* (66), *MUC6*, *MUC5B* and *MUC5AC* were identified as CDX2 targets. To our knowledge, this has not previously been demonstrated and suggests *CDX2* as a locus control gene of the 11p15 mucin gene cluster (67–69) that includes *MUC2*, *MUC6*, *MUC5B* and *MUC5AC*. In this respect, a recent study of 295 gastric adenocarcinomas has shown that *GPA33*, *MUC6* and *MUC2* were among the 10 most abundant differentially regulated genes in 295 analyzed microsatellite instable (MSI) colorectal adenocarcinomas (70). Although *CDX2* profiling was not evident in this study it is stimulating to assume that MSI has impacted on *CDX2* expression in the tumors studied. Furthermore, *GPA33* is

expressed in cancer of the gastrointestinal tract, namely in over 95% of human colon cancers and has thus been suggested as an attractive novel colorectal carcinoma therapeutic target (71,72). Of particular notice, *GPA33* gene product has been shown to be post translationally modified by the *GALNT1* encoded GalNAc-transferase 1 enzyme (73,74), that we here show is inversely influenced by CDX2 (Figure 5B). Our finding that *LDLR* is a major direct CDX2 target gene is surprising since *LDLR* has been primarily found to be regulated by sterol regulated element-binding proteins (SREBPs) (75,76). We speculate that CDX2 plays novel roles in the co-regulated appearance of these novel gene targets.

In summary, we hereby describe a novel approach for generating an inducible 'non leaky' isogenic knockout-rescue system. The principle is based on precise genome integration of defined copies of third generation Tet-On elements, which enables for an 'all or none' context for studying the mechanism by which a given gene works in a cell. We exemplify the utility of the system by establishing a colonic CDX2 PrIITE cell system that uncovered novel molecular interactions governed by CDX2.

MATERIALS & CORRESPONDENCE

All reagents and cell lines used in the study are available upon request for research purposes under a material transfer agreement, except for TRE3G and TET3G containing plasmids, due to restrictions over distribution of plasmids containing the TRE3G promoter or TET3G ORF. The RNA-seq and ChIP-seq data shown has been uploaded at the NIH GEO server: http://www.ncbi.nlm.nih.gov/geo/ and is accessible using the following accession number GSE97273. Plasmids described are available from Addgene (https://www.addgene.org/). All correspondence should be addressed to Eric P. Bennett.

NOTE ADDED IN PROOF

Extended experiments conducted during proofs of this manuscript revealed that the GPA33 promoter described in Figure 6 possess bi-directional transcriptional activity, likely mediated through presence of multiple LINEs contained within the 1KBp GPA33 promoter analyzed. LINEs have been shown to mediate bi-directional expression (Trinklein, N.D., Aldred, S.F., Hartman, S.J., Schroeder, D.I., Otillar, R.P. and Myers, R.M. (2004) An abundance of bidirectional promoters in the human genome. Genome Res.14:62–66; Core, L.J., Waterfall, J.J. and Core, J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science, 322:1845–1848) potentially affecting 4% of all human genes (Speek, M. (2001) Antisense promoter of human L1 retrotransposon drives transcription of adjacent cellular genes. Mol. Cell. Biol. 2001, 1973–1985; Criscione, S.W., Theodosakis, N., Micevic, G., Cornish, T.C., Burns, K.H., Neretti, N., and Rodić, N. (2016) Genome-wide characterization of humanL1 antisense promoter-driven transcripts. BMC Genomics, 17:463).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Camilla Andersen, Claire Katrine Beier Holden, Asli Silahtaroglu, Elisabeth Larsen, Mads Bak (Departments of Odontology and Cellular and Molecular Medicine, Copenhagen Center for Glycomics, University of Copenhagen) and Katja Dahlgaard, (Department of Science and Environment, Roskilde University, Denmark) for excellent technical assistance. We thank Kim Madsen at CLC Bioscience for guidance and access to CLC Workbench.

Author contributions: R.P. planned and performed all ZFN targeting, IHC, FACS and live imaging experiments. L.H., M.C and C.M. planned and performed RNA-seq and ChIPSeq experiments. S.L. and J.D. performed all *in vitro* promoter analysis. J.B.H. performed Dox induction and FACS experiments. J.T.T., R.A. and L.D. contributed with writing parts of the manuscript. E.P.B. designed and planned all experiments and wrote the manuscript.

FUNDING

University of Copenhagen Excellence Programme for Interdisciplinary Research [CDO2016]; The Danish Council for Independent Research [4004-00140B], Novo Nordisk Foundation, Danish National Research Foundation [DNRF107], FEDER—Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020—Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020; Portuguese funds through FCT—Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Inovação [POCI-01-0145-FEDER-007274, POCI-01-0145-FEDER-016390]. Funding for open access charge: Danish National Research Foundation [DNRF107].

Conflict of interest statement. None declared.

REFERENCES

1. McClintock,B. (1950) the origin and behavior of mutable loci in maize. *Proc. Natl. Acad. Sci. U.S.A.*, **36**, 344–355.
2. Goth,R. and Rajewsky,M.F. (1972) Ethylation of nucleic acids by ethylnitrosourea-l-14C fetal and adult Rat1. *Cancer Res.*, **32**, 1501–1505.
3. Hinton,C. and Whittinghill,M. (1950) The distribution of X-ray induced crossovers from curly inversion heterozygotes of Drosophila melanogaster females. *Proc. Natl. Acad. Sci. U.S.A.*, **36**, 552–558.
4. Capecchi,M.R. (1980) High efficiency transformation by direct microinjection of DNA into cultured mammalian cells. *Cell*, **22**, 479–488.
5. Mali,P., Yang,L., Esvelt,K.M., Aach,J., Guell,M., DiCarlo,J.E., Norville,J.E. and Church,G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
6. Cong,L., Ran,F.A., Cox,D., Lin,S., Barretto,R., Habib,N., Hsu,P.D., Wu,X., Jiang,W., Marraffini,L.a. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
7. Geurts,A.M., Cost,G.J., Freyvert,Y., Zeitler,B., Jeffrey,C., Choi,V.M., Jenkins,S.S., Wood,A., Cui,X., Meng,X. *et al.* (2010) Knockout rats produced using designed zinc finger nucleases. *Science*, **325**, 2009–2011.
8. Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by

double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806–811.
9. Liu,J., Deutsch,U., Fung,I. and Lobe,C.G. (2014) Conditional and inducible transgene expression in endothelial and hematopoietic cells using Cre/loxP and tetracycline-off systems. *Exp. Ther. Med.*, **8**, 1351–1356.
10. Kuhn,R., Schwenk,F., Aguet,M. and Rajewsky,K. (1995) Inducible gene targeting in mice. *Science*, **269**, 1427–1428.
11. Braselmann,S., Graninger,P. and Busslinger,M. (1993) A selective transcriptional induction system for mammalian cells based on Gal4-estrogen receptor fusion proteins. *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 1657–1661.
12. Kennedy,M.J., Hughes,R.M., Peteya,L.A., Schwartz,J.W., Ehlers,M.D. and Tucker,C.L. (2010) Rapid blue-light – mediated induction of protein interactions in living cells. *Nat. Methods*, **7**, 12–16.
13. No,D., Yao,T.P. and Evans,R.M. (1996) Ecdysone-inducible gene expression in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.*, **93**, 3346–3351.
14. Unger,B. and Hillen,W. (1984) Nucleotide sequence of the repressor gene of the RAI tetracycline resistance determinant: structural and functional comparison with three related Tet repressor genes Bernhard. *Nucleic Acids Res.*, **12**, 7693–7703.
15. Postle,K., Nguyen,T.T. and Bertrand,K.P. (1984) Nucleic acids research. *Nucleic Acids Res.*, **12**, 4849–4863.
16. Gossen,M. and Bujard,H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. U.S.A.*, **89**, 5547–5551.
17. Gossen,M., Freundlieb,S., Bender,G., Müller,G., Hillen,W. and Bujard,H. (1995) Transcriptional activation by tetracyclines in mammalian cells. *Science*, **268**, 1766–1769.
18. Das,A.T., Zhou,X., Metz,S.W., Vink,M.A. and Berkhout,B. (2016) Selecting the optimal Tet-On system for doxycycline-inducible gene expression in transiently transfected and stably transduced mammalian cells. *Biotechnol. J.*, **11**, 71–79.
19. Urlinger,S., Baron,U., Thellmann,M., Hasan,M.T., Bujard,H. and Hillen,W. (2000) Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 7963–7968.
20. Freundlieb,S., Schirra-Müller,C. and Bujard,H. (1999) A tetracycline controlled activation/repression system with increased potential for gene transfer into mammalian cells. *J. Gene Med.*, **1**, 4–12.
21. Loew,R., Heinz,N., Hampf,M., Bujard,H. and Gossen,M. (2010) Improved Tet-responsive promoters with minimized background expression. *BMC Biotechnol.*, **10**, 81.
22. Heinz,N., Schambach,A., Galla,M., Maetzig,T., Baum,C., Loew,R. and Schiedlmeier,B. (2011) Retroviral and transposon-based tet-regulated all-in-one vectors with reduced background expression and improved dynamic range. *Hum. Gene Ther.*, **22**, 166–176.
23. Das,A.T., Liliane,T. and Berkhout,B. (2016) Tet-On systems for doxycycline-inducible gene expression. *Curr. Gene Ther.*, **16**, 156–167.
24. Moullan,N., Mouchiroud,L., Houtkooper,R.H., Auwerx,J., Moullan,N., Mouchiroud,L., Wang,X., Ryu,D., Williams,E.G. and Mottis,A. (2015) Tetracyclines disturb mitochondrial function across eukaryotic models: a call for caution in biomedical article tetracyclines disturb mitochondrial function across eukaryotic models: a call for caution in biomedical research. *Cell Rep.*, **10**, 1681–1691.
25. Bensinger,S.J., Graeber,T.G. and Christofk,H.R. (2013) Doxycycline alters metabolism and proliferation of human cell lines. *PLoS One*, **8**, 8–14.
26. Brüning,A., Brem,G.J., Vogel,M. and Mylonas,I. (2013) Tetracyclines cause cell stress-dependent ATF4 activation and mTOR inhibition. *Exp. Cell Res.*, **320**, 281–289.
27. Zhou,X., Vink,M., Klaver,B., Berkhout,B. and Das,A.T. (2006) Optimization of the Tet-On system for regulated gene expression through viral evolution. *Gene Ther.*, **13**, 1382–1390.
28. Suh,E., Chen,L., Taylor,J. and Traber,P.G. (1994) A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Mol. Cell. Biol.*, **14**, 7340–7351.
29. Duda,K., Lonowski,L.a., Kofoed-Nielsen,M., Ibarra,A., Delay,C.M., Kang,Q., Yang,Z., Pruett-Miller,S.M., Bennett,E.P., Wandall,H.H. *et al.* (2014) High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs. *Nucleic Acids Res.*, doi:10.1093/nar/gku251.
30. Maresca,M., Lin,V.G., Guo,N. and Yang,Y. (2013) Obligate Ligation-Gated Recombination (ObLiGaRe): custom-designed nucleases mediated targeted integration through nonhomologous end joining. *Genome Res.*, doi:10.1101/gr.145441.112.
31. Wandall,H.H., Rumjantseva,V., Sørensen,A.L.T., Patel-Hett,S., Josefsson,E.C., Bennett,E.P., Italiono,J.E., Henrik,C., H,H.J. and Hoffmeister,K.M. (2012) The origin and function of platelet glycosyltransferases Glycosyltransferase assays Immunoblot analysis Flow cytometric analysis of platelet glycosyltransferases. *Blood*, **120**, 626–635.
32. Yang,Z., Steentoft,C., Hauge,C., Hansen,L., Thomsen,a. L., Niola,F., Vester-Christensen,M.B., Frodin,M., Clausen,H., Wandall,H.H. *et al.* (2015) Fast and sensitive detection of indels induced by precise gene targeting. *Nucleic Acids Res.*, **43**, e59.
33. Lonowski,L.A., Narimatsu,Y., Riaz,A., Delay,C.E., Yang,Z., Niola,F., Duda,K., Ober,E.A., Clausen,H., Wandall,H.H. *et al.* (2017) Genome editing using FACS enrichment of nuclease-expressing cells and indel detection by amplicon analysis. *Nat. Protoc.*, **12**, 581–603.
34. Coskun,M., Troelsen,J.T. and Nielsen,O.H. (2011) The role of CDX2 in intestinal homeostasis and inflammation. *Biochim. Biophys. Acta*, **1812**, 283–289.
35. Ji,H., Jiang,H., Ma,W., Johnson,D.S., Myers,R.M. and Wong,W.H. (2008) An integrated system CisGenome for analyzing ChIP-chip and ChIP-seq data. *Nat. Biotechnol.*, **26**, 1293–1300.
36. Gao,N., White,P. and Kaestner,K.H. (2010) Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Dev. Cell*, **16**, 588–599.
37. Mesquita,P., Jonckheere,N., Almeida,R., Ducourouble,M.P., Serpa,J., Silva,E., Pigny,P., Santos Silva,F., Reis,C., Silberg,D. *et al.* (2003) Human MUC2 mucin gene is transcriptionally regulated by Cdx homeodomain proteins in gastrointestinal carcinoma cell lines. *J. Biol. Chem.*, **278**, 51549–51556.
38. Boyd,M., Hansen,M., Jensen,T.G.K., Perearnau,A., Olsen,A.K., Bram,L.L., Bak,M., Tommerup,N., Olsen,J. and Troelsen,J.T. (2010) Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). *J. Biol. Chem.*, **285**, 25115–25125.
39. Bennett,E.P., Mandel,U., Clausen,H., Gerken,T.a, Fritz,T.a and Tabak,L.a (2012) Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*, **22**, 736–756.
40. Salari,K., Spulak,M.E., Cuff,J., Forster,A.D., Giacomini,C.P. and Huang,S. (2012) CDX2 is an amplified lineage-survival oncogene in colorectal cancer Keyan. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 3196–3205.
41. Rutzky,L.P., Kaye,C.I., Siciliano,M.J., Chao,M. and Kahan,B.D. (1980) Longitudinal karyotype and genetic signature analysis of cultured human colon adenocarcinoma cell lines LS180 and LS174T. *Cancer Res.*, **40**, 1443–1448.
42. Göttke,M.S., Keller,K., Belley,A., Garcia,R.-M., Hollingsworth,M.A., Mack,D.R. and Chadee,K. (1998) Functional heterogeneity of colonic adenocarcinoma mucins for inhibition of Entamoeba histolytica adherence to target cells. *J. Eukaryot. Microbiol.*, **45**, 17S–23S.
43. Bu,X.D., Li,N., Tian,X.Q. and Huang,P.L. (2011) Caco-2 and LS174T cell lines provide different models for studying mucin expression in colon cancer. *Tissue Cell*, **43**, 201–206.
44. Heath,J.K., White,S.J., Johnstone,C.N., Catimel,B., Simpson,R.J., Moritz,R.L., Tu,G.F., Ji,H., Whitehead,R.H., Groenen,L.C. *et al.* (1997) The human A33 antigen is a transmembrane glycoprotein and a novel member of the immunoglobulin superfamily. *Proc. Natl. Acad. Sci. U.S.A.*, **94**, 469–474.
45. Pinto,R., Barros,R., Pereira-Castro,I., Mesquita,P., da Costa,L.T., Bennett,E.P., Almeida,R. and David,L. (2015) CDX2 homeoprotein is involved in the regulation of ST6GalNAc-I gene in intestinal metaplasia. *Lab. Invest.*, **95**, 1–10.
46. Sakuma,K., Aoki,M. and Kannagi,R. (2012) Transcription factors c-Myc and CDX2 mediate E-selectin ligand expression in colon cancer cells undergoing EGF / bFGF-induced epithelial—mesenchymal transition. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 7776–7781.

47. Isshiki,S., Kudo,T., Nishihara,S., Ikehara,Y., Togayachi,A., Furuya,A., Shitara,K., Kubota,T., Watanabe,M., Kitajima,M. *et al.* (2003) Lewis type 1 antigen synthase (β3Gal-T5) is transcriptionally regulated by homeoproteins. *J. Biol. Chem.*, **278**, 36611–36620.
48. Mitchelmore,C., Troelsen,J.T., Spodsberg,N., Sjöström,H. and Norén,O. (2000) Interaction between the homeodomain proteins Cdx2 and HNF1alpha mediates expression of the lactase-phlorizin hydrolase gene. *Biochem. J.*, **346**, 529–535.
49. Hansen,L., Lind-thomsen,A., Joshi,H.J., Pedersen,N.B., Have,C.T., Kong,Y., Wang,S., Sparso,T., Grarup,N., Vester-christensen,M.B. *et al.* (2015) A glycogene mutation map for discovery of diseases of glycosylation. *Glycobiology*, **25**, 211–224.
50. Rizzino,A. (2009) Sox2 and Oct-3/4: a versatile pair of master regulators that orchestrate the self-renewal and pluripotency of embryonic stem cells. *Wiley Interdiscip. Rev. Syst. Biol. Med.*, **1**, 228–236.
51. Yang,Z., Wang,S., Halim,A., Schulz,M.A., Frodin,M., Rahman,S.H., Vester-Christensen,M.B., Behrens,C., Kristensen,C., Vakhrushev,S.Y. *et al.* (2015) Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.*, doi:10.1038/nbt.3280.
52. Bahr,S., Cortner,L., Ladley,S. and Borgschulte,T. (2013) Evaluating the effect of chromosomal context on zinc finger nuclease efficiency. *BMC Proc.*, **7**, P3.
53. Mansour,W.Y., Schumacher,S., Rosskopf,R., Rhein,T., Schmidt-Petersen,F., Gatzemeier,F., Haag,F., Borgmann,K., Willers,H. and Dahm-Daphi,J. (2008) Hierarchy of nonhomologous end-joining, single-strand annealing and gene conversion at site-directed DNA double-strand breaks. *Nucleic Acids Res.*, **36**, 4088–4098.
54. Cristea,S., Freyvert,Y., Santiago,Y., Holmes,M.C., Urnov,F.D., Gregory,P.D. and Cost,G.J. (2013) In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnol. Bioeng.*, **110**, 871–880.
55. Yang,Z., Wang,S., Halim,A., Schulz,M.A., Frodin,M., Rahman,S.H., Vester-christensen,M.B., Behrens,C., Kristensen,C., Vakhrushev,S.Y. *et al.* (2015) Engineered CHO cells for production of diverse , homogeneous glycoproteins. *Nat. Biotechnol.*, **33**, 842–844.
56. Inui,M., Miyado,M., Igarashi,M., Tamano,M., Kubo,A., Yamashita,S., Asahara,H., Fukami,M. and Takada,S. (2014) Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. *Sci. Rep.*, **4**, 5396.
57. Lee,J.S., Grav,L.M., Pedersen,L.E., Lee,G.M. and Kildegaard,H.F. (2016) Accelerated homology-directed targeted integration of transgenes in Chinese hamster ovary cells via CRISPR/Cas9 and fluorescent enrichment. *Biotechnol. Bioeng.*, **9999**, 1–17.
58. He,X., Tan,C., Wang,F., Wang,Y., Zhou,R., Cui,D., You,W., Zhao,H., Ren,J. and Feng,B. (2016) Knock-in of large reporter genes in human cells via CRISPR/Cas9-induced homology-dependent and independent DNA repair. *Nucleic Acids Res.*, **44**, e85.
59. Dow,L.E., Fisher,J., O'Rourke,K.P., Muley,A., Kastenhuber,E.R., Livshits,G., Tschaharganeh,D.F., Socci,N.D. and Lowe,S.W. (2015) Inducible in vivo genome editing with CRISPR-Cas9. *Nat. Biotechnol.*, **33**, 390–394.
60. Wang,J., Exline,C.M., DeClercq,J.J., Llewellyn,G.N., Hayward,S.B., Li,P.W.-L., Shivak,D.A., Surosky,R.T., Gregory,P.D., Holmes,M.C. *et al.* (2015) Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.*, **33**, 1256–1263.
61. Cai,Y., Laustsen,A., Zhou,Y., Sun,C., Anderson,M.V., Li,S., Uldbjerg,N., Luo,Y., Jakobsen,M.R. and Mikkelsen,J.G. (2016) Targeted, homology-driven gene insertion in stem cells by ZFN-loaded ' all-in-one ' lentiviral vectors. *Elife*, e12213.
62. Coluccio,A., Miselli,F., Lombardo,A., Marconi,A., Malagoli Tagliazucchi,G., Gonçalves,M.a, Pincelli,C., Maruggi,G., Del Rio,M., Naldini,L. *et al.* (2013) Targeted gene addition in human epithelial stem cells by zinc-finger nuclease-mediated homologous recombination. *Mol. Ther.*, **21**, 1695–1704.
63. Moehle,E.a, Rock,J.M., Lee,Y.-L., Jouvenot,Y., DeKelver,R.C., Gregory,P.D., Urnov,F.D. and Holmes,M.C. (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 3055–3060.
64. Yamamoto,Y., Bliss,J. and Gerbi,S.A. (2015) Whole organism genome editing: targeted large DNA insertion via ObLiGaRe nonhomologous end-joining in vivo capture. *G3 (Bethesda).*, **5**, 1843–1847.
65. Yang,Z., Wang,S., Halim,A., Schulz,M.A., Frodin,M., Rahman,S.H., Vester-Christensen,M.B., Behrens,C., Kristensen,C., Vakhrushev,S.Y. *et al.* (2015) Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.*, **33**, 2014–2017.
66. Brown,M.S., Goldstein,J.L., Arnold,W. and Perdue,S. (1974) Expression of the familial hypercholesterolemia gene in heterozygotes: mechanism for a dominant disorder in man. *Science*, **185**, 61–63.
67. Van Cong,N., Aubert,J.P., Gross,M.S., Porchet,N., Degand,P. and Frézal,J. (1990) Assignment of human tracheobronchial mucin gene(s) to 11p15 and a tracheobronchial mucin-related sequence to chromosome 13. *Hum. Genet.*, **86**, 167–172.
68. Gerard,C., Eddy,R.L. and Shows,T.B. (1990) The core polypeptide of cystic fibrosis tracheal mucin contains a tandem repeat structure: evidence for a common mucin in airway and gastrointestinal tissue. *J. Clin. Invest.*, **86**, 1921–1927.
69. Velcich,A., Palumbo,L., Selleri,L., Evans,G. and Augenlicht,L. (1997) Organization and regulatory aspects of the human intestinal mucin gene (MUC2) locus. *J. Biol. Chem.*, **272**, 7968–7976.
70. Wheeler,D.a., Gibbs,R.a., Muzny,D.M., Bainbridge,M.N., Chang,K., Dinh,H.H., Drummond,J.a., Fowler,G., Kovar,C.L., Lewis,L.R. *et al.* (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature*, **487**, 330–337.
71. Sakamoto,J., Kojima,H., Kato,J., Hamashima,H. and Suzuki,H. (2000) Organ-specific expression of the intestinal epithelium-related antigen A33, a cell surface target for antibody-based imaging and treatment in gastrointestinal cancer. *Cancer Chemother. Pharmacol.*, **46**(Suppl), S27–S32.
72. Ciprotti,M., Chong,G., Gan,H.K., Chan,A., Murone,C., MacGregor,D., Lee,F.-T., Johns,T.G., Heath,J.K., Ernst,M. *et al.* (2014) Quantitative intratumoural microdistribution and kinetics of (131)I-huA33 antibody in patients with colorectal carcinoma. *EJNMMI Res.*, **4**, 1–10.
73. Steentoft,C., Vakhrushev,S.Y., Joshi,H.J., Kong,Y., Vester-Christensen,M.B., Schjoldager,K.T.-B.G., Lavrsen,K., Dabelsteen,S., Pedersen,N.B., Marcos-Silva,L. *et al.* (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.*, **32**, 1478–1488.
74. Kong,Y., Joshi,H.J., Schjoldager,K.T.-B.G., Madsen,T.D., Gerken,T.a, Vester-Christensen,M.B., Wandall,H.H., Bennett,E.P., Levery,S.B., Vakhrushev,S.Y. *et al.* (2014) Probing polypeptide GalNAc-transferase isoform substrate specificities by in vitro analysis. *Glycobiology*, **0**, 1–11.
75. Horton,J.D., Goldstein,J.L. and Brown,M.S. (2002) SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin. Invest.*, **109**, 1125–1131.
76. Shimano,H. (2001) Sterol regulatory element-binding proteins (SREBPs): transcriptional regulators of lipid synthetic genes. *Prog. Lipid Res.*, **40**, 439–452.

Supplementary material:

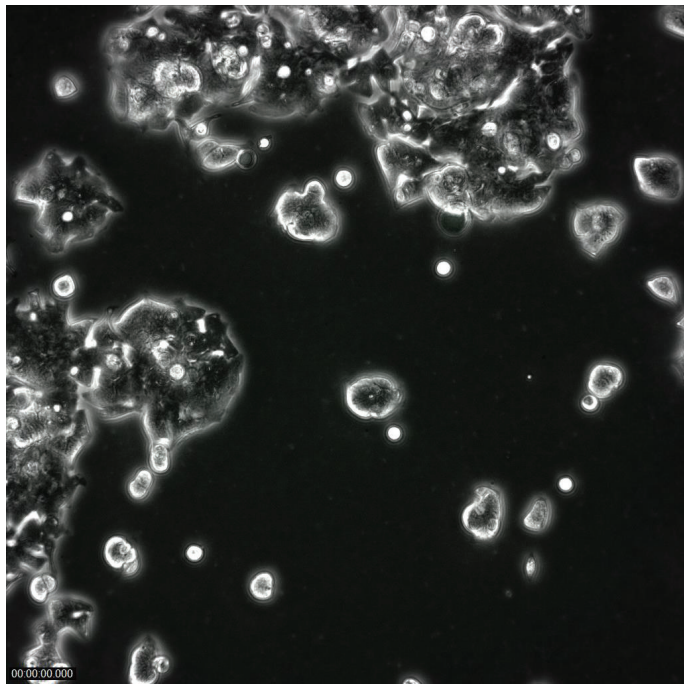
Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression.

Rita Pinto, Lars Hansen, John Hintze, Raquel Almeida, Sylvester Larsen, Mehmet Coskun, Johanne Davidsen, Cathy Mitchelmore, Leonor David, Jesper Thorvald Troelsen and Eric Paul Bennett

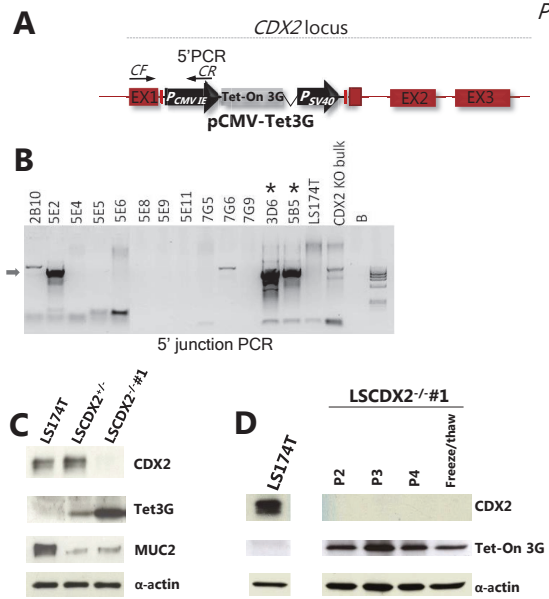
Supplementary material contains:

Supplementary video material
Supplementary Figures 1-11
Supplementary Tables I-IV

LS174T T2EYFP PrIITE Cells

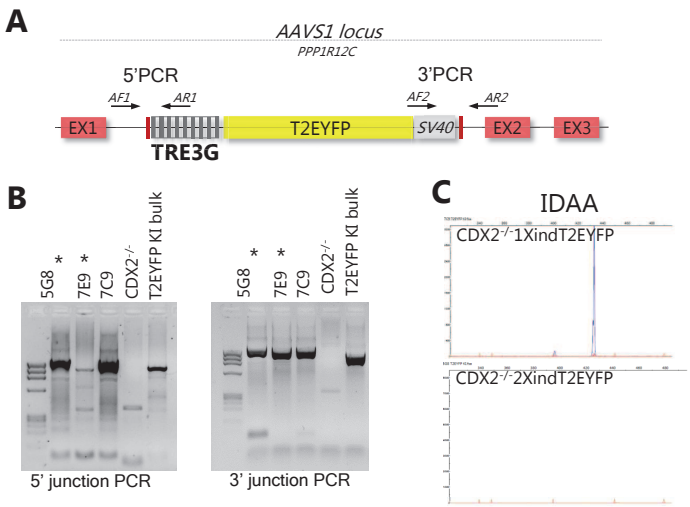


Supplementary video material
24h Time-lapse video showing T2EYFP expression in CDX2$^{-/-}$2XindT2EYFP clone upon induction with 0.5µg/mL of doxycycline.

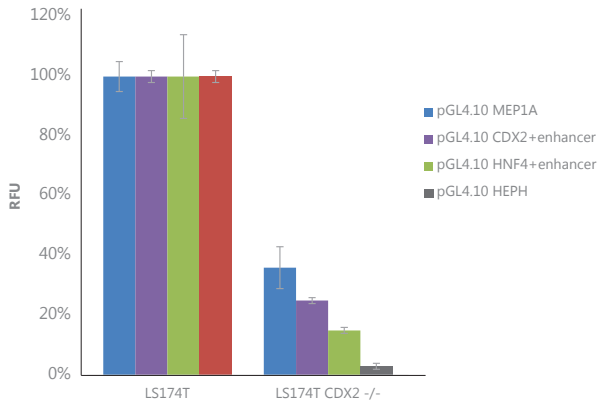


Supplementary Figure 1

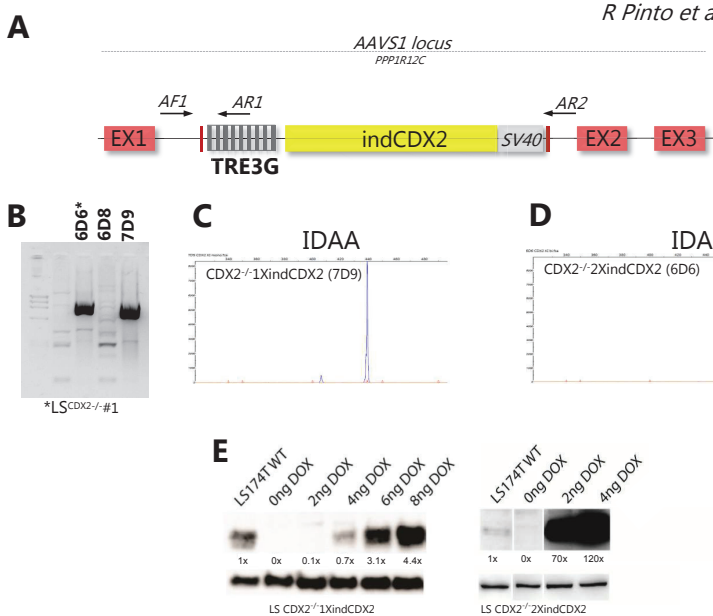
Target integration of tetracycline responsive transactivator (TET3G) donor elements into *CDX2* exon 1. **A.** Schematic diagram of the pCMV-Tet3G targeting construct, with the regulatory elements controlling constitutive Tet3G expression, targeted to *CDX2* exon1. Forward and reverse primers used for determining proper target integration at the *CDX2* locus are shown (CF and CR respectively). **B.** Correct integration was detected by successful amplification of a 1.3 kb product (5'-junction PCR). Furthermore, amplification with primers flanking the *CDX2* ZFN target site would detect wild type (wt) allele presence, while lack of successful amplification would indicate correct targeting due to amplification settings chosen, not allowing for amplification across the integrated donor target sequence (data not shown). Asterix indicates clones included in this study. LS174T and CDX2KO bulk represent untransfected and 2 days post electroporation transfected cells respectively. No positive signal could be detected on bulk by target integration using homologous recombination and 800bp homology flanking donor arms on donor construct, data not shown. **C.** Western blot detection of CDX2, Tet3G and MUC2 in LS174Twt, $CDX2^{+/-}$ or $CDX2^{-/-}$ cells. Monoclonals against CDX2, Tet3G, MUC2 and α-actin (control) show, that CDX2 protein is only expressed in wt and $CDX2^{+/-}$ cells and absent in $CDX2^{-/-}$ cells. 104 clones were analyzed in total and copy number dependent Tet3G expression was detected in the respective PrIITE cells and MUC2 protein was clearly decrease in expression in the cells shown. α-actin control is shown below. **D.** Western blot analysis of LS174Twt cells and bi-allelic targeted 3D6 clone ($CDX2^{-/-}$) in the latter case after prolonged passage (P2, P3 and P4) and freeze/thawing cycle. Stable Tet3G expression was observed and absence of *CDX2* maintained.



Supplementary Figure 2
Target integration of T2EYFP "proof of principle" reporter donor elements to *AAVS1* intron 1. **A.** Schematic illustration of *p*AAVS1-TRE3G-T2EYFP targeted to AAVS1 intron 1 locus. Position of the genetic primers used for determining correct targeting are shown with arrows above the Tet3G responsive elements/TRE3G promoter and SV40 3'UTR, AF1/2 and AR1/2. **B.** Junction PCR showing correct integration of donor construct at the *AAVS1* target site. Asterix indicates clones included in this study. CDX2$^{-/-}$ and T2EYFP KI bulk represent untransfected and 2day post transfection controls. **C.** Detection of wild type allele by IDAA assay[50]. Presence of detected amplicon/peak represents untargeted wild type allele while lack of peak represents bi-allelic target integration of *p*AAVS1-TRE3G-T2EYFP donor construct.

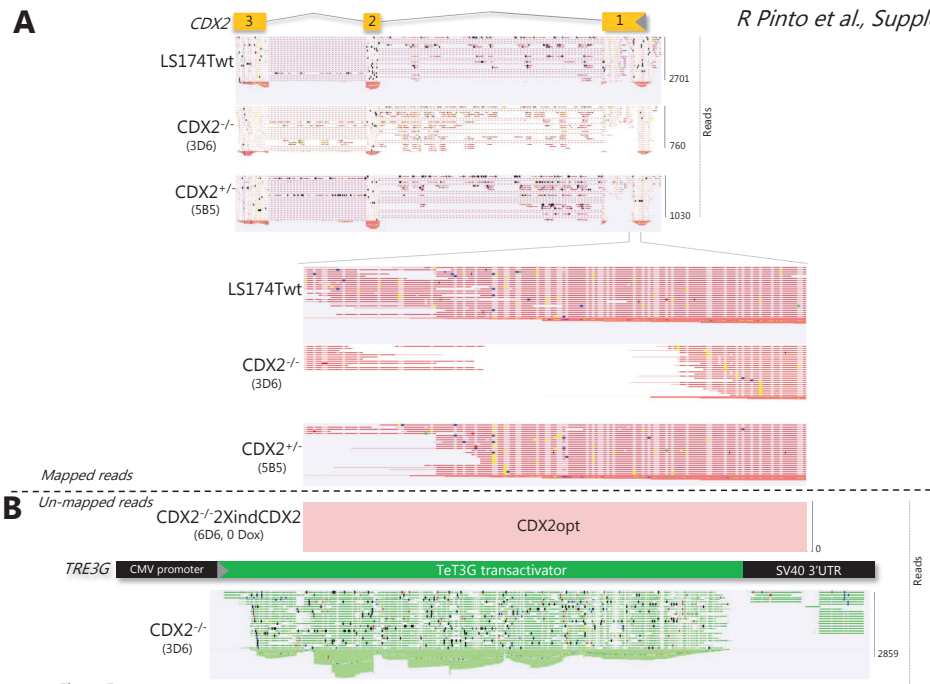


Supplementary Figure 3
Validating the lack of functional CDX2 expression in the CDX2$^{-/-}$ 3D6 clone. Using known CDX2 dependent enhancer elements[30], promoter analysis was conducted in LS174T cells and 3D6 cells and the relative fluorescence units normalized to levels observed in LS174Twt cells. Most prominent >95% repressed expression was observed for the HEPH reporter, clearly demonstrating the lack of functional CDX2 in the 3D6 cells.



Supplementary Figure 4

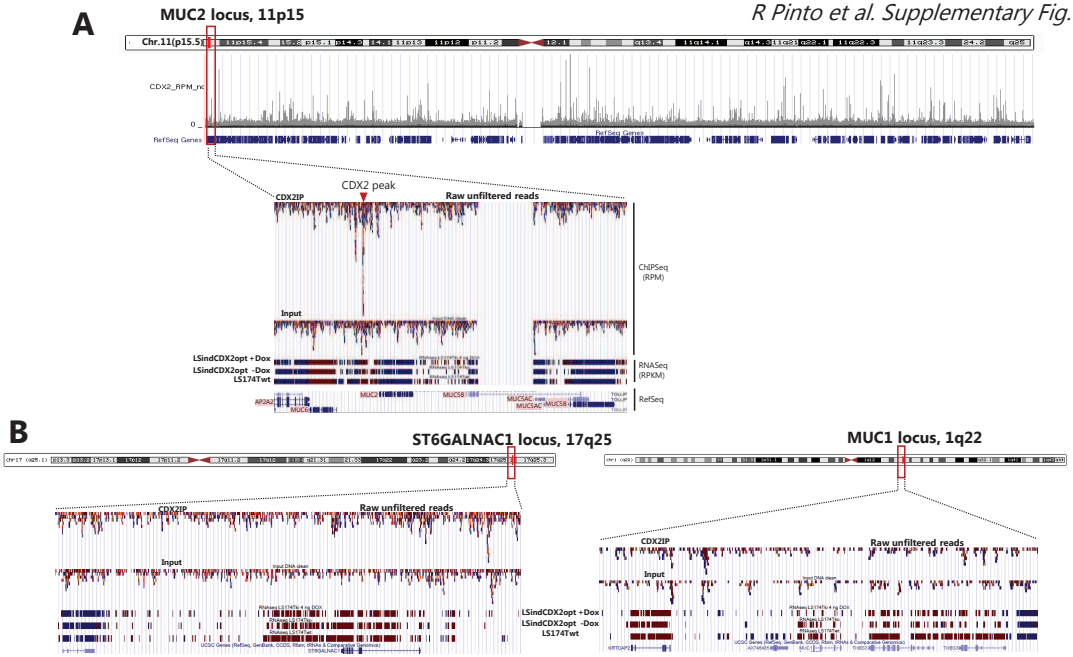
Target integration of Tet3G inducible codon optimized CDX2 donor elements to *AAVS1* intron 1. **A.** Schematic illustration of the *p*AAVS1-TRE3G-CDX2opt donor construct targeted to *AAVS1* intron 1. Position of the genetic primers used for determining correct targeting are shown with arrows above the Tet3G responsive element/TRE3G promoter and flanking the SV40 3'UTR, AF1, AR1/2. **B.** Junction PCR showing correct target integration in CDX2$^{-/-}$ cells (LS$^{CDX2-/-}$#1). Clone 6D6 was used for further analyses in this study, indicated by an asterix. **C.** Detection of wild type *AAVS1* allele by IDAA assay[50]. Presence of detected amplicon/peak represents untargeted wild type allele presence and thus integration of only one copy of *p*AAVS1-TRE3G-CDX2opt donor construct (CDX2$^{-/-}$1XindCDX2 (7D9)). **D.** Lack of peak represents bi-allelic target integration and thus integration of 2 copies of *p*AAVS1-TRE3G-CDX2opt donor construct (CDX2$^{-/-}$2XindCDX2 (6D6)). **E.** Western blot analysis for CDX2 expression in LS174Twt and PrIITE cells (LS CDX2$^{-/-}$1XindCDX2 or LS CDX2$^{-/-2}$XindCDX2) with or without variable Dox concentrations. Equal amounts of total protein (10µg) were loaded on the primary CDX2-88 anti-body probed Westerns shown. Lower panels illustrate loading control probings with anti GAPDH (Fitzgerald Industries International, USA, cat#10R-G109a) or anti Vinculin (Abcam, EPR8185) respectively. For clarity results were quantified by densitometry using ImageJ 1.51j8 software (https://imagej.nih.gov/ij/download.html) and presented as fold difference relative to CDX2 levels detected in LS174Twt cells.



R Pinto et al., Supplementary Fig. 5

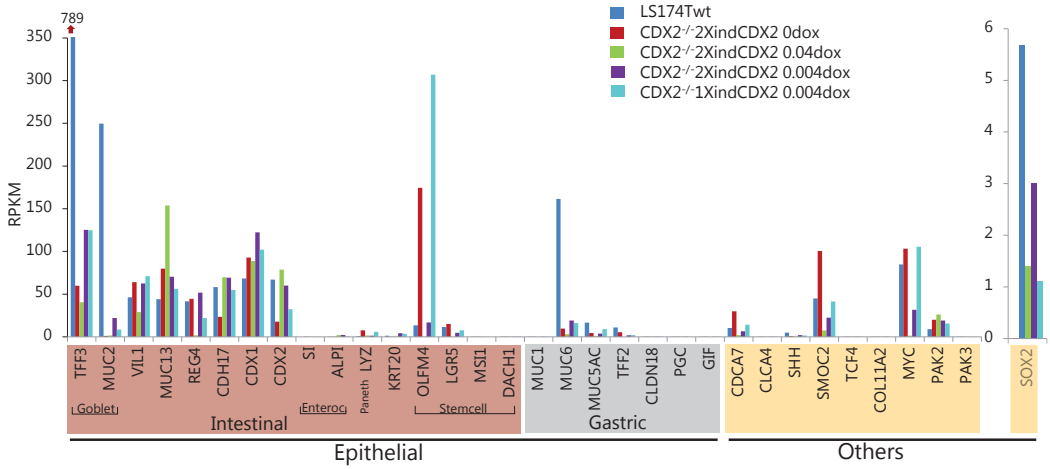
Supplementary Figure 5

Detailed insight of the LS-PrIITE cell derived RNA-seq transcriptome data set. **A.** Upper panel depicts the *CDX2* locus, the direction of transcription is indicated by an arrow. Sequence fragments/read densities for LS174Twt, *CDX2*$^{-/-}$ (3D6) and *CDX2*$^{+/-}$ (5B5) are shown across the locus as red lines and below each RNA-seq panel the exon read densities are indicated with red profiles in the ligth blue shaded panels. The total number of locus reads is indicated to the right of each panel. An expanded view of the pCMV-Tet3G donor integration targeted site within *CDX2* exon 1 is shown below the three upper RNA-seq panels. Note the lack of reads in the CDX2$^{-/-}$ 3D6 clone at the exact sites of biallelic pCMV-Tet3G integration and that in spite of this, *CDX2* locus specific transcripts are detectable but at ≈25% abundance relative to LS174Twt levels. **B.** Detailed insight into un-mapped reads of RNAseq sequences that could not be matched to sequences in the hg19 database. Since codon optimized *CDX2* and *Tet3G* transactivator encoded donor integration elements were used, transcripts from these respective transcription units were not identified in the hg19 mapped reads data set. Upper empty light blue panel depicts the complete absence of inducible codon optimized CDX2 transcripts in uninduced CDX2$^{-/-}$2XindCDX2 6D6 cells. Lower panel depicts the abundant reads identified for the constitutively expressed Tet3G transactivator in CDX2$^{-/-}$ 3D6 cells that found the basis for most of the cells included in this study. Number of reads are indicated to the right of the panels.



Supplementary Figure 6

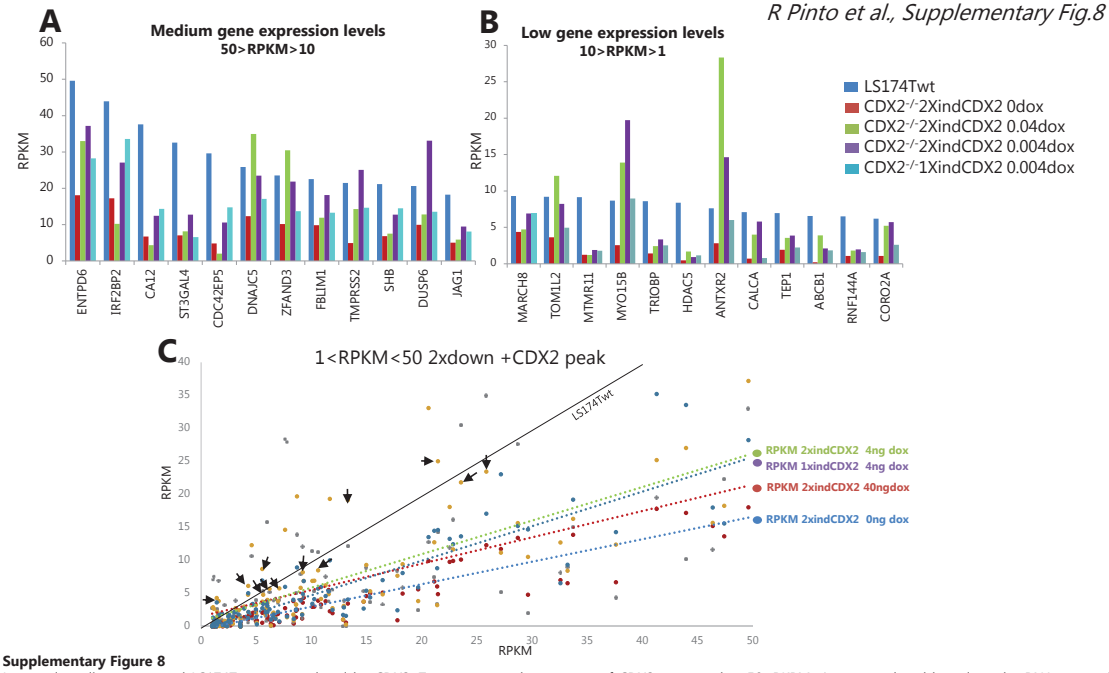
RNA-seq and ChIPSeq identified primary CDX2 target genes. **A.** Panel shows chromosome 11 coverage of CDX2IP (CDX2 ChIPseq) reads. Coverage was determined as described in **Fig.6**. Major CDX2 binding sites (peaks) are shown with indication of the *MUC2* 11p15 mucin gene locus framed in red box. Expanded display of the 11p15 mucin gene locus, showing detailed identification of CDX2 ChIP-Seq (CDX2IP) and control (Input) peaks in both LS174Twt samples. The major *MUC2* CDX2 binding element is indicated by red triangle (CDX2 peak). Lower RNA-seq panels display the density of individual reads in CDX2$^{-/-}$2XindCDX2 cells with or without Dox induction (4ng or 0ng respectively) and LS174Twt cells. RefSeq gene panel displays the genomic organization of genes within the 11p15 locus. Notably, the major peak (CDX2 enhancer element) was positioned in the intergenic *MUC6* and *MUC2* region. The deleted region shown in panel is due to an unannotated region in the hg19 genome reference data base used for the analysis. **B.** Chromosome 17 and 1 diplays with zoom in panels shown below, displaying absence of CDX2 peaks in the *ST6GALNAC1* and *MUC1* respective loci. The former locus has previously been suggested to be regulated by CDX2[34], whereas the latter locus has not been linked to CDX2 regulation.



Supplementary Figure 7

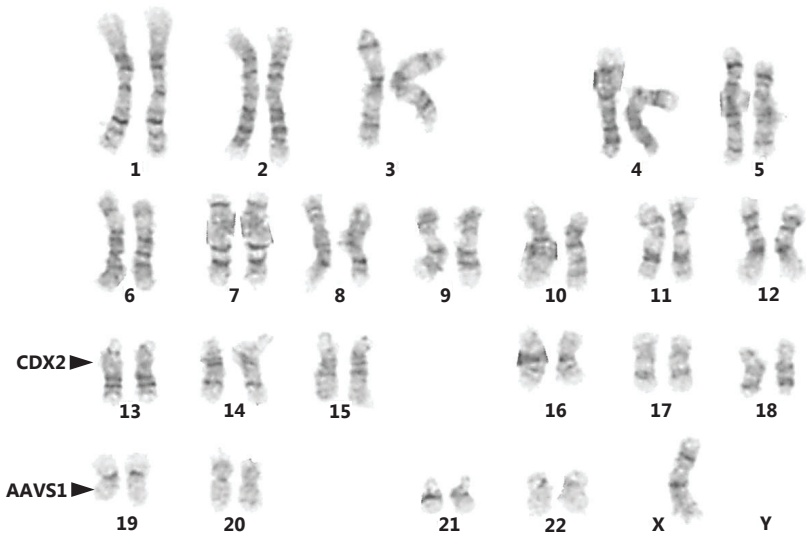
Lineage and stemness gene expression profiles for LS-CDX2 PrIITE cells. RPKMs for selected epithelial marker genes and others are shown for LS174Twt (dark blue), CDX2$^{-/-}$2XindCDX2 without induction (red), CDX2$^{-/-}$2XindCDX2 with 40ng/ml Dox induction (green), CDX2$^{-/-}$2XindCDX2 with 4ng/ml Dox induction (purple) and CDX2$^{-/-}$1XindCDX2 with 4ng/ml Dox induction (light blue). RNA-seq expression patterns of intestinal and gastric specific marker genes in both LS174Twt and PrIITE cells, with or without induction were highly similar. Clearly the cells used and generated in this study displayed a mixed goblet, epithelial and stemcell expression profile, but notably lacked complete expression of the SI and ALPI enterocyte markers. Of notice, the gastric specific marker gene MUC6 was also found to be expressed in LS174Twt cells, in contrast *SOX2* was only found to be expressed in the CDX2$^{-/-}$ PrIITE cells generated.

R Pinto et al., Supplementary Fig.8



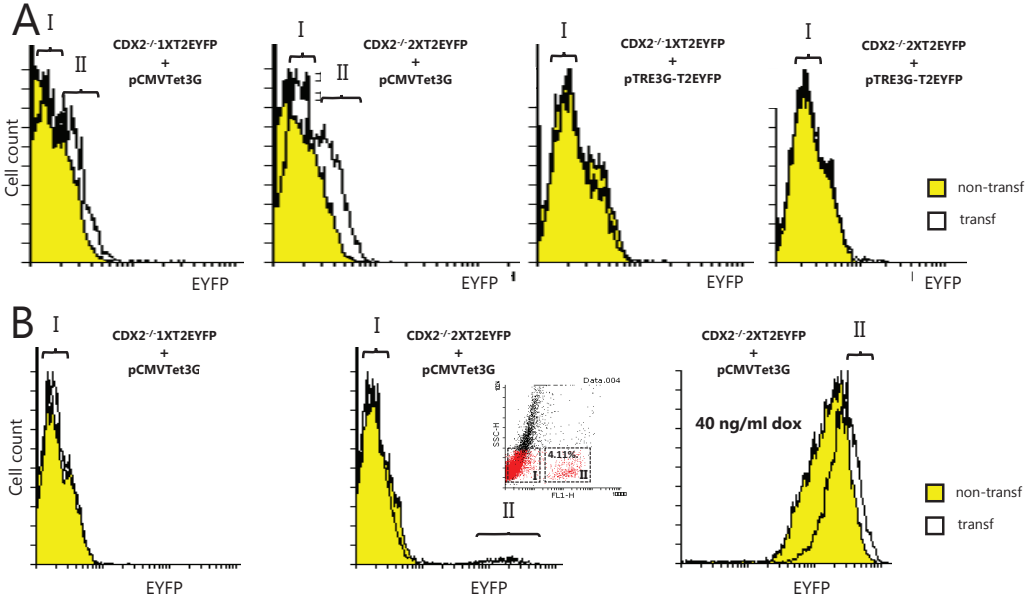
Supplementary Figure 8

Low and medium expressed LS174T genes regulated by CDX2. Targets genes downstream of CDX2 expressed at 50>RKPM>1 were analyzed based on the RNA-seq and ChIPSeq criterions 4X reduced and 2X increased RPKM in the respective uninduced and induced state and with CDX2 ChIPSeq peak in locus. The reversibility profiles for the genes in the respective PrIITE cells is similar as for the genes profiles displayed for highly expressed genes (RKPMs>50) shown in **Figure 7**, showing that 2XindCDX2 cells responded better than 1XindCDX2 cells. List of CDX2 regulated genes with 1>RKPMs >0 is given **Supplementary Table III. A.** Profiles for genes expressed with RPKMs between 10 and 50. **B.** Profiles for genes expressed with RPKMs between 10 and 1. **C.** Global Dox induction reversibility of the PrIITE cells relative to LS174Twt cells for low and medium level expressed genes (1<RPKM<50). Data analysis conditions as described in **Fig.7**, RNA-seq data (RPKM) from wildtype LS174T was plotted against RPKM values from uninduced CDX2$^{-/-}$2XindCDX2 cells without CDX2 expression (blue dotted line, 2XindCDX 0ng/ml DOX) and under different conditions of CDX2/Dox induction (ng/ml) (green dotted line, 2XindCDX2 4ng/ml DOX; purple dotted line, 1XindCDX2 4ng/ml DOX; red dotted line, 2XindCDX2 40ng/ml DOX). The regression line for each data set is plotted. The dark blue solid line represents the expression levels in LS174Twt cells. Representative examples for genes displaying near full reversibility are indicated with arrow heads. The data shows, that the induction profiles for low and medium expressed genes are similar to the global profiles shown in **Fig.7**.



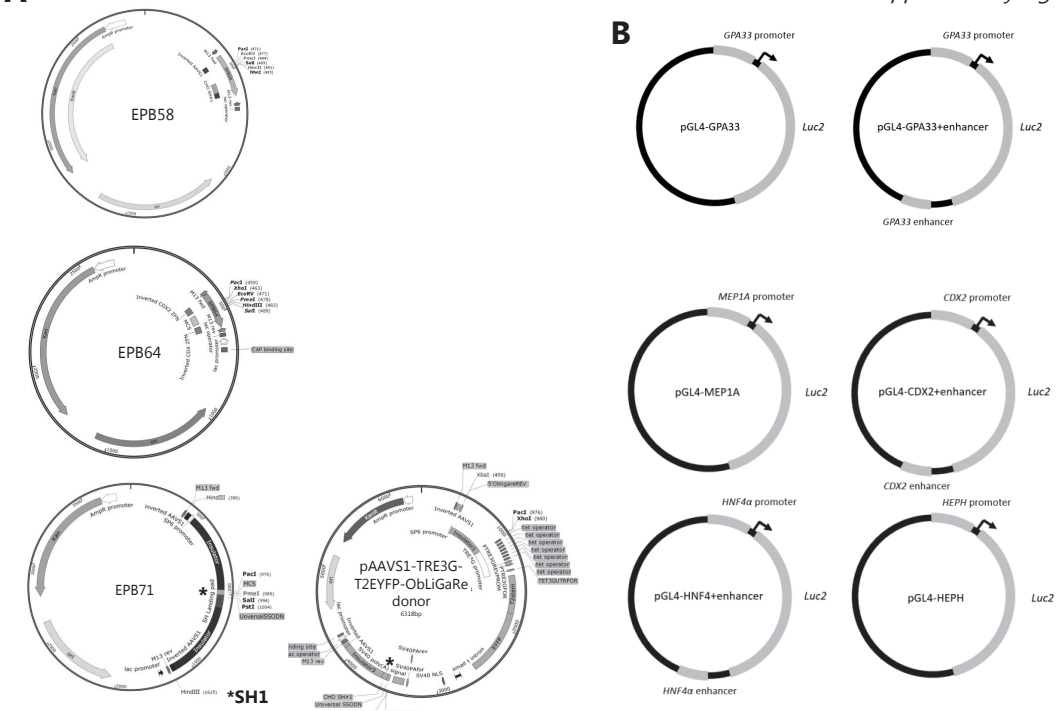
Supplemantary Figure 9
Karyogram of LS174T. The LS174T cell line used in this study was karyotyped using standard Giemsa-banding and karyotyping methods with slight modifications. Cells were treated with colcemid (0.004% w/v) for 1.5 hrs and trypsinized for 3-5 minutes before harvesting. All 25 cells analyzed from both the wild type and the modified cell line showed a 45,X karyotype with occasional 16q abnormality which was in agreement with the karyotype of the published LS174T cell line originating from a female adenocarcinoma (LP Rutzky et al., Cancer Research, 1980) displaying a near normal karyotype. The loci targeted in this study are indicated on the relevant chromosomes.

R Pinto et al. Supplementary Fig.10



Supplementary Figure 10

Effect of increased expression and stable random integration on "leaky expression" of Tet3G inducible elements. **A.** LS174T PrIITE cells with defined precise targeted copies of TET3G and pTRE3G T2EYFP used as proof of principle, were transfected with either pCMVTet3G (left two panels) or pTRE3G-T2EYFP (right two panels) and the transiently transfected cell pools, 2 days post transfection, were analyzed by FACS for detection of potential "leaky" T2EYFP expression in the absence of Dox induction. Overlays of FACS analysis of non-transfected and transfected cells clearly showed, that transient overexpression of pCMVTet3G transactivator gave rise to two distinct population of cells; (I) a non fluorescent population identical to non transfected cells and (II) a population with increased leaky expression of T2EYFP distinct from the non-transfected cells. Transient over expression of the inducible pTRE3G-T2EYFP plasmid did not seem to give rise to leaky T2EYFP expression. **B.** Stable cell pools of the pCMVTet3G transfected cells analyzed in panel A, were obtained after up to 45 days G418 selection, followed by FACS analysis. Left panel shows that CDX2-/-1XT2EYFP cells do not display leaky expression as a consequence of increased expression of randomly integrated pCMVTE3G expression. Middle panel clearly shows that a 4.11% sub-population of cells (see side scatter insert for clarity) maintain substantial leaky T2EYFP expression in CDX2-/-2XT2EYFP cells at levels similar to 24h 40ng/ml Dox induction levels (right panel). Note, that the induction potential is increased in a stable randomly integrated pCMVTET3G sub-population of cells II). We speculated that these cells represent the cells with leaky expression in the non-induced state.



Supplementary Figure 11

Vector maps of the plasmid constructs used in this study. All plasmids used available from Addgene (Cambridge, USA). **A.** Plasmid maps for the target integration constructs used for establishing the various PrIITE cells. EPB58 (Addgene ID#90016) was used for insertion of TRE3G codon optimized CDX2 establishing *pAAVS1-TRE3G-CDX2opt ObLiGaRe-donor vector*. EPB64 (Addgene ID#90017) represents was used for inserting Tet3G transactivator into donor vector targeting the CDX2 exon 1 locus. EPB71 (Addgene ID#90018) used for establishing the T2EYFP AAVS1 donor integration vector pAAVS1-TRE3G-T2EYFP ObLiGaRe-donor vector shown. **B.** Vector maps for reporter plasmids used for promoter analysis of the various enhancer elements tested *in vitro*. Position of the CHO Safe Harbor #1 (SH1) "landing pad" is marked by an asterix.

Supplementary information:

Supplementary Table I. CDX2 as an activator*

Gene	LS174Twt CDX2$^{+/+}$ (rpkm)	CDX2$^{-/-}$2XindCDX2 0,00µg/ml dox	CDX2$^{-/-}$2XindCDX2 0,04µg/ml dox	CDX2$^{-/-}$2XindCDX2 0,004µg/ml dox	CDX2$^{+/-}$1XindCDX2 0,004µg/ml dox	>4 fold down	>2 fold up
TFF3	789,43	59,77	40,37	125,21	124,79	13,21	2,09
GPA33°	317,18	21,29	44,15	109,84	95,37	14,90	5,16
MUC2°	249,49	1,15	1,58	22,00	8,52	217,63	19,19
ETV4	164,05	22,13	45,55	89,86	64,74	7,41	4,06
ETHE1	110,69	25,75	40,56	71,57	65,55	4,30	2,78
LDLR	78,65	12,33	27,87	45,34	20,50	6,38	3,68
CDC42EP5	29,63	4,84	2,03	10,56	14,73	6,13	2,18
TMPRSS2	21,46	4,91	14,26	25,05	14,64	4,37	5,10
DFNB31	17,79	0,97	8,44	7,71	2,47	18,42	7,98
IRS2	11,67	2,32	5,38	19,35	10,03	5,02	8,32
TRIOBP	8,60	1,41	2,42	3,34	2,54	6,10	2,37
CALCA	7,09	0,69	4,00	5,80	0,79	10,22	8,35
ABCB1	6,57	0,22	3,89	2,11	1,83	30,38	9,78
CORO2A	6,20	1,08	5,23	5,72	2,61	5,77	5,31
PTK2B°	6,01	1,14	15,76	4,04	2,21	5,29	3,56
SLC6A20	5,53	0,59	2,67	8,73	6,97	9,42	14,86
SHH°	4,97	0,84	0,62	2,06	1,43	5,94	2,46
EGF	3,53	0,16	0,18	0,35	0,24	21,96	2,17

VPS37D	3,53	0,13	1,68	1,44	0,32	26,62	10,90
PDE9A°	2,66	0,17	0,33	0,80	0,64	15,69	4,71
THSD4	2,22	0,16	0,42	0,63	0,46	13,88	3,92
ATXN1	2,13	0,24	3,30	0,87	0,44	8,78	3,58
PTPRO	1,94	0,03	0,03	0,14	0,23	59,17	4,12
ARL14	1,58	0,36	6,91	3,87	1,74	4,42	10,78
B4GALNT3	1,48	0,26	1,16	1,58	0,75	5,79	6,15
LTBP1	1,46	0,10	0,59	0,42	0,28	13,88	3,96
KRT20	1,41	0,02	0,52	4,33	3,67	84,16	258,23
RBP3	1,20	0,04	0,01	0,26	0,10	27,75	6,06
CRB2	1,19	0,18	0,32	1,11	0,43	6,63	6,19
AOAH	1,15	0,01	0,63	1,15	0,22	133,41	132,89
CNBD1	1,01	0,01	0,02	0,19	0,15	83,96	15,82

* 4x↓in KO, 2x ↑+Dox (0.004µg/ml), CDX2 ChIP-seq peak in locus

° CDX2 targets previously identified in CaCo2 cells[1]

Supplementary TableII. CDX2 as a repressor*

Name	LS174Twt CDX2$^{+/+}$ (rpkm)	CDX2$^{-/-}$2XindCDX2 0,00µg/ml dox	CDX2$^{-/-}$2XindCDX2 0,04µg/ml dox	CDX2$^{-/-}$2XindCDX2 0,004µg/ml dox	CDX2$^{-/-}$1XindCDX2 0,004µg/ml dox
C4orf45	0	0,042776646	0	0	0,0242754
CD84	0	0,013292883	0,0094436	0,0075729	0,0037718
CFHR2	0	0,025168911	0	0	0,0285663
HOXA4	0	0,035210542	0,0166763	0	0,0199817
HOXC5	0	0,340791232	0,177545	0,0388293	0
IL17A	0	0,016544599	0	0	0
LIX1	0	0,022439014	0	0	0,0594251
MGAT4C	0	0,006539742	0,0123893	0	0
MORC1	0	0,008171203	0,00774	0	0,0092742
OR10V1	0	0,029516707	0	0	0
PTPRC	0	0,003615424	0,0205479	0	0,0041034
RPS6KA6	0	0,063854136	0,0302424	0	0,0362367
SLAMF6	0	0,302411884	0	0,0893313	0,3305203
TACR3	0	0,011852181	0,0112268	0,0067521	0
TEKT2	0	0,031098492	0,0441863	0	0

*Genes not expressed in LS174Twt, 4x ↑ in KO, 2x ↓+DOX (0.004µg/ml), CDX2 ChIP seq peak in locus

Supplementary Table III. CDX2 as an activator*

Name	LS174Twt CDX2$^{+/+}$ (rpkm)	CDX2$^{-/-}$2XindCDX2 0,00μg/ml dox	CDX2$^{-/-}$2XindCDX2 0,04μg/ml dox	CDX2$^{-/-}$2XindCDX2 0,004μg/ml dox	CDX2$^{-/-}$1XindCDX2 0,004μg dox
ANGPT1	0,7900189	5,5272512	3,9543323	4,6432396	1,9913809
C1orf21	0,7876576	3,2099251	1,9464836	3,0472531	3,4268473
AMACR	0,6386579	5,4593419	0,9981764	2,0324455	4,5895585
C9orf156	0,5126537	2,5889785	2,1962671	2,0630245	2,2875263
CABLES1	0,4026463	3,0876587	4,4746709	5,255985	3,3103341

*Low expressed genes 0<RPKM<1, 4x↓ in KO, 2x ↑+Dox induction (0,004μg/ml), CDX2 ChIP-seq peak in locus

Supplementary Table IV

	Sequence (5'->3')
Junction PCR	*CDX2* locus: CF: GGG CTC TCT GCT TGT CAC CTA CCA GG CR: CAC GCC CAT TGA TGT ACT GCC *AAVS1* locus: AF1: GCC CTC TAA CGC TGC CGT CTC AR2: TAT AGG CGC CCA CCG TAC ACG CC AF2: CACACCTCCCCCTGAACCTGA AR2: CGT AAG CAA ACC TTA GAG GTT CTG G
IDAA analysis	*CDX2* locus : CDX2S: AGC TGA CCG GCA GCA AAA TTG GAC GTG AGC ATG TAC CCT AGC TC CDX2AS: CTG CGC GCT GTC CAA GTT CGC TG *AAVS1* locus: AAVS1S: AGC TGA CCG GCA GCA AAA TTG CCT TAC CTC TCT AGT CTG TGC TAG AR2: CGT AAG CAA ACC TTA GAG GTT CTG G FAMF: FAM-AGC TGA CCG GCA GCA AAA TTG

Paper II

The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2

Davidsen J, Larsen S, Coskun M, Gögenur I, Dahlgaard K, Bennett EP, and Troelsen JT.

PLoS One 2018, 13(7)

A recurrent fusion of the *VTI1A* and *TCF7L2* genes has been identified in approximately 3 % of colorectal cancer patients. This study investigates the characteristics of the resulting fusion protein, VTI1A-TCF7L2 (in the paper referred to as VTI1A-TCF4), as well as the transcriptional regulation of the protein. As wild-type TCF7L2 is crucial in stem cell niche maintenance and proliferation in the intestinal crypts, the fusion protein with truncated TCF7L2 may play an undesirable role in colon cancer development.

This study demonstrates that the VTI1A-TCF7L2 fusion protein is most likely a dominant negative regulator of Wnt signaling, and results indicate that CDX2 may activate transcription of the fusion protein, resulting in the dominant negative regulator being expressed throughout the entire axis of the crypt. The results of this study may help to uncover one possible pathway in colon cancer development, as well as highlight the effect of altered promoter regions of fusion genes.

PLOS ONE

RESEARCH ARTICLE

The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2

Johanne Davidsen[1,2], Sylvester Larsen[1,3], Mehmet Coskun[4], Ismail Gögenur[2], Katja Dahlgaard[1], Eric Paul Bennett[5], Jesper Thorvald Troelsen[1]*

1 Department of Science and Environment, Roskilde University, Roskilde, Denmark, **2** Department of Surgery, Zealand University Hospital, Roskilde, Denmark, **3** Department of Clinical Immunology, Naestved Hospital, Naestved, Denmark, **4** Department of Gastroenterology, Herlev Hospital, Herlev, Denmark, **5** Copenhagen Center for Glycomics, Department of Odontology, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

* troelsen@ruc.dk

Check for updates

OPEN ACCESS

Citation: Davidsen J, Larsen S, Coskun M, Gögenur I, Dahlgaard K, Bennett EP, et al. (2018) The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2. PLoS ONE 13(7): e0200215. https://doi.org/10.1371/journal.pone.0200215

Editor: Masaru Katoh, National Cancer Center, JAPAN

Received: February 19, 2018

Accepted: June 21, 2018

Published: July 5, 2018

Copyright: © 2018 Davidsen et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Sundhed og Sygdom, Det Frie Forskningsråd (4004-00140B) (Dr Jesper Thorvald Troelsen); University of Copenhagen Excellence Programme for Interdiciplinary Research (CDO2016) (Dr Eric Paul Bennett); and Danmarks Grundforskningsfond

Abstract

Sequencing of primary colorectal tumors has identified a gene fusion in approximately 3% of colorectal cancer patients of the *VTI1A* and *TCF7L2* genes, encoding a VTI1A-TCF4 fusion protein containing a truncated TCF4. As dysregulation of the Wnt signaling pathway is associated with colorectal cancer development and progression, the functional properties and transcriptional regulation of the VTI1A-TCF4 fusion protein may also play a role in these processes. Functional characteristics of the VTI1A-TCF4 fusion protein in Wnt signaling were analyzed in NCI-H508 and LS174T colon cancer cell lines. The NCI-H508 cell line, containing the *VTI1A-TCF7L2* fusion gene, showed no active Wnt signaling, and overexpression of the VTI1A-TCF4 fusion protein in LS174T cells along with a Wnt signaling luciferase reporter plasmid showed inhibition of activity. The transcriptional regulation of the *VTI1A-TCF4* fusion gene was investigated in LS174T cells where the activity of the *VTI1A* promoter was compared to that of the *TCF7L2* promoter, and the transcription factor CDX2 was analyzed for gene regulatory activity of the *VTI1A* promoter through luciferase reporter gene assay using colon cancer cell lines as a model. Transfection of LS174T cells showed that the *VTI1A* promoter is highly active compared to the *TCF7L2* promoter, and that CDX2 activates transcription of *VTI1A*. These results suggest that the VTI1A-TCF4 fusion protein is a dominant negative regulator of the Wnt signaling pathway, and that transcription of *VTI1A* is activated by CDX2.

Introduction

Colorectal cancer is one of the most commonly diagnosed types of cancer in the western world and a leading cause of cancer-related death. The mechanisms behind the development of

(DNRF107) (Dr Eric Paul Bennett). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

sporadic colorectal cancer are, despite considerable research, not fully understood [1,2]. Yet, disruptions of the canonical Wnt signaling pathway are known to play a major role in cancer initiation as well as progression and it is estimated that 80–90% of all colorectal cancer tumors harbor mutant Adenomatous Polyposis Coli (APC), an essential scaffold protein in the Wnt signaling pathway [1,3,4].

The central signal transducer in the canonical Wnt signaling pathway is β-catenin. In the absence of Wnt glycoprotein ligands, β-catenin is phosphorylated and subsequently degraded in proteasomes. The interaction between the GSK3 and CK1α kinases and β-catenin is facilitated by the scaffold proteins APC and axin, and combined, the kinases and the scaffold proteins form the degradation complex [5]. Upon binding of secreted Wnt glycoprotein to transmembrane co-receptors, the canonical Wnt signaling pathway is activated which leads to accumulation of intracellular β-catenin. β-catenin then enters the nucleus where it associates with members of the T-cell factor/Lymphoid enhancer-binding factor (TCF/LEF) family of transcription factors and activates transcription of Wnt target genes by displacing the Groucho co-repressor bound to TCF/LEF proteins. If APC or other members of the degradation complex have loss-of-function mutations, β-catenin will not be degraded, resulting in a constitutively active Wnt signaling pathway [1,5,6]. The main binding partner of β-catenin in the colon is TCF4, and studies have shown that TCF4 plays an important part in maintaining the proliferative cells in the colonic crypts, and that natural downregulation of TCF4 expression in colonic epithelial cells migrating up the crypt may induce differentiation [6–8].

Dysregulation of members of the TCF/LEF family of transcription factors has been observed in both colon cancer cell lines, and colon cancer tumors [1,9], and there is indication that the TCF/LEF family of transcription factors work in distinct and opposing roles to maintain the equilibrium between epithelial cell proliferation and terminal differentiation in normal colonic epithelium. LEF1 is specifically expressed in early stages of B-cell differentiation but has also been shown to be expressed in colon cancer tumors [3,10,11]. TCF1 expression is largely restricted to T-lymphocytes in adult tissue but expression has also been detected in colorectal cancer cell lines [12]. In adult mice with a dominant mutated *APC* gene, conditional knockout of TCF4 significantly enhances colon tumor formation, indicating that TCF4 is a tumor suppressor [3]. However, the expression of TCF4 in colon tumors has been shown to correlate to lower survival, indicating oncogenic properties of TCF4 [1]. Thus, the role of TCF4 in colon cancer is not yet fully understood and it may function as both a tumor suppressor and an oncogene [1,3]. With disruptions in the Wnt signaling pathway found in the majority of colon cancer tumors it is nonetheless to be expected that dysregulation of TCF4 has a part to play in colon cancer initiation and/or progression.

Through genomic sequencing, Bass et al. (2011) has identified a recurrent fusion of the *Vps-ten-interacting-1a* (*VTI1A)* and the *T-cell factor 7-like 2* (*TCF7L2)* genes in 3% of primary colorectal adenocarcinomas. The fusion fuses the first three exons of the *VTI1A* gene with the fourth exon of the neighboring gene, *TCF7L2*, thereby omitting the first three exons of the *TCF7L2* gene. The *VTI1A* gene encodes a v-SNARE protein that mediates vesicle transport from late endosomes to the trans-Golgi network, and the *TCF7L2* gene encodes the TCF4 transcription factor. The fusion of the *VTI1A* and *TCF7L2* genes results in a fusion protein, VTI1A-TCF4, in which the N-terminal β-catenin binding domain of TCF4 is lacking, while it still contains the DNA binding domain and the transcription repression domain [13].

Nome et al. (2014) has also detected the VTI1A-TCF7L2 fusion transcript in colorectal cancer tissue samples showing that the genomic rearrangement is functional. They have further shown that 42% of colorectal cancers express the fusion transcript and have detected seven different splice variants of the transcript. The fusion transcripts were also detected in 28% of normal colonic mucosal samples and 5 normal tissue samples from different anatomical sites of

the body. The fact that fusions involving *TCF7L2* are highly detectable in colorectal cancers, normal colonic tissue and other normal tissue types, may indicate that these fusions are neither specific to cancer nor to the colon or rectum [14]. However, the similar fusion transcripts observed in individual cancer cell lines induced by genomic rearrangements may still have oncogenic potential. In this study we investigate the transcript generated by the deletion in the NCI-H508 colon cancer cell line, a cell line harboring the gene fusion [13].

The *VTI1A* promoter most likely regulates the transcription of the VTI1A-TCF4 fusion protein. ChIP-seq data indicates that the *VTI1A* promoter contains binding sites for caudal type homeobox 2 (CDX2) [15,16], which is expressed in differentiating cells of the colonic crypts [17,18], and not in the proliferating cells, where TCF4 is expressed [8]. This means that the VTI1A-TCF4 fusion protein may be expressed throughout the entire colonic crypt and not only at the base where expression of TCF4 is normally seen. Further, research indicates that dysregulation of CDX2 in the colon plays a role in the development and progression of colorectal cancer [19–21].

The VTI1A-TCF4 fusion protein has been shown to play a critical role in anchorage-independent growth of the colon cancer cell line NCI-H508, indicating that the VTI1A-TCF4 protein has functional properties [13]. Other members of the TCF/LEF family of transcription factors naturally occur as truncated isoforms with no β-catenin binding domains, creating proteins with dominant negative properties [22]. The functional properties of the VTI1A-TCF4 fusion protein have not yet been investigated, but research suggests that it might also exhibit dominant negative properties as it also lacks the β-catenin binding domains [23,24]. The aim of this study was to determine the functional properties of the VTI1A-TCF4 fusion protein and to investigate the level of transcriptional activity from the *VTI1A-TCF7L2* fusion gene in colon cancer cells compared to that of wildtype *TCF4*. Additionally, it aims to explore the role of CDX2 in regulation of transcription of the *VTI1A-TCF7L2* fusion gene.

Materials and methods

Cloning of *VTI1A* and *TCF7L2* promoter-reporter constructs

The *VTI1A* promoter region (735 bp, from position -765 to -30, NM_145206) was amplified from human genomic DNA using the VTI1A-735 primer set and cloned with the In-Fusion HD cloning system (Clontech, #639649) into the *HindIII* site of the modified firefly luciferase reporter pGL4.10 vector, resulting in the pGL4.10-VTI1A-735 construct. Plasmids for deletion analysis were created using the same reverse primer but different forward primers. The promoter region for *TCF7L2* (1405 bp, from position -1545 to -140, NM_030756) was amplified from human genomic DNA using the TCF7L2-1405 primer set and cloned into the pGL4.10 vector using the *Xho1* and *HindIII* sites, resulting in the pGL4.10-TCF7L2-1405 construct. The plasmids were sequence verified (Beckman Coulter Genomics). Primer sequences can be found in S1 Table.

Cloning of VTI1A-TCF4 expression plasmid

NCI-H508 cells were seeded at 500.000 cells per well in a 6-well plate and incubated at 37˚C and 5% CO_2 for 48 hours. The cells were harvested and total RNA was purified with Total RNA Kit I (Omega Bio-tek, #R5834-02). RNA to cDNA synthesis was performed using First Strand cDNA Synthesis kit (Thermo Scientific, #K1612) according to manufactures protocol. The cDNA was amplified with VTI1A-TCF4 sequence specific primers and gel-purified *VTI1A-TCF7L2* sequence was cloned into *NheI* digested pLVX-CherryPicker2 vector (Clontech, #632581) using the In-Fusion cloning system (Clontech, #639649) resulting in the VTI1A-TCF4 fusion protein expression plasmid. The plasmid was sequence verified

(Beckman Coulter Genomics). Primer sequences can be found in S1 Table. Protein expression from the VTI1A-TCF4 expression plasmid was verified by Western Blot, see S1 Fig. LS174T cells were seeded at 600.000 cells per well in a 6-well plate with or without 1.2 μg VTI1A-TCF4 expression plasmid, according to transfection protocol below, and incubated at 37˚C and 5% CO_2 for 48 hours. The cells were then lysed in ice cold 1XLDS buffer in RIPA buffer. Each sample was heated to 82˚C for 3 minutes and separated on 4–12% BisTris SDS PAGE with 1XMOPS SDS running buffer. Proteins were transferred to PVDF membrane 0.45 μm (Invitrolon), and the membrane was incubated with TCF4/TCF7L2 (C48H11) Rabbit mAb (Cell Signaling) 1:1000 overnight at 4˚C and subsequently incubated for 2 hours with Pierce 332260 Goat Anti-Rabbit IgG, HRP-linked Antibody 1:5000 followed by detection by Pierce's Dura detection reagent.

Cell culture

LS174T, NCI-H508, and Caco-2 cell lines were used in this study. LS174T and Caco-2 cells were grown in DMEM 4.5 g/L Glucose with UltraGlutamine (Lonza, #BE12-604F-U1) containing 10% heat-inactivated Fetal Bovine Serum and 100 U/ml penicillin and streptomycin. The NCI-H508 cells were cultured in RMPI 1640 medium with L-Glutamine (Lonza, #BE12-702F) added 10% Fetal Bovine Serum and 100 U/ml penicillin and streptomycin. All cells were grown at 37˚C and 5% CO_2 in T75 flasks and passaged twice a week. An LS174T cell line with both *CDX2* alleles genetically knocked out, obtained from Pinto et al. (2017), was also used. These cells were cultured as the wild type LS174T cell line.

Transfection

For transfections, LS174T cells and NCI-H508 cells were seeded at 100.000 cells per well in 24-well plates. Caco-2 cells were seeded at 50.000 cells per well. Cell transfections were done using 2 μM Polyethyleneimine (Alfa Aesar #43896). Four replicas were made and each transfection was carried out at least twice. The luciferase reporter plasmid pGL4.10-VTI1A was used for the transfections, and the pCMV-lacZ plasmid was used to determine transfection efficiency control. The TOPFlash and FOPFlash reporter system was used to determine Wnt signaling [25]. For co-transfections, plasmids expressing CDX2, or the VTI1A-TCF7L2 fusion protein were added the transfection. To normalize to a total amount of 0.3 μg DNA per well pBluescript SK+ plasmid was added. 48 hours after transfection the cells were harvested, and the luciferase and β-galactosidase activities were analyzed using the Dual Light assay kit (Invitrogen, #T1004) on a GloMax 96 Microplate Luminometer with dual injectors (Promega). The activity for both luciferase and β-galactosidase were measured for 5 seconds, and each luciferase activity was normalized to β-galactosidase activity. The activity is stated in relative fluorescence units (RFU).

Chromatin immunoprecipitation

Confluent LS174T cells were cross-linked by formaldehyde treatment and then sonicated. Immunoprecipitation with CDX2 and HA was done as described in Coskun et al. 2012 [26]. Verification of immunoprecipitation enrichment was performed with quantitative PCR (qPCR) using a Stratagene MX3005P (Agilent Technologies) cycler with Quantitect SYBR Green PCR Mastermix (Qiagen, #204141) according to manufactures protocol. The primer set CDX2 ChIP was used and can be found in S1 Table. Quantification of the ChIP-DNA was done using the delta-delta method described by Livak and Schmittgen, 2001 [27].

Statistical methods

Values are presented as means with standard deviations, and groups were compared using two-sided Student's t-test. P-values lower than 0.05 were considered significant. * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$, **** indicates $p<0.0001$.

Results

The VTI1A-TCF4 fusion protein has dominant negative properties

The level of Wnt signaling in LS174T, NCI-H508, and Caco-2 cell lines was investigated using the TOPFlash/FOPFlash reporter system [25]. The TOPFlash reporter plasmid contains seven TCF binding sites upstream a luciferase reporter gene that will be transcribed during active Wnt signaling. The FOPFlash reporter plasmid works as a negative control and contains mutated TCF binding sites upstream a luciferase reporter gene and can therefore not be activated by binding of TCF4 during Wnt signaling [25]. Results showed that the transcriptional activity of the TOPFlash reporter plasmid in LS174T cells was more than 500-fold higher compared to the activity of the FOPFlash reporter plasmid ($p<0.0001$, Fig 1b). A ~150-fold higher TOPFlash transcriptional activity was seen in Caco-2 cells compared to FOPFlash activity ($p<0.0001$). Both cell lines have loss-of-function mutations in regulators of the Wnt signaling pathway. The LS174T cell line has mutated *β-catenin*, making CK1α unable to bind [28], and the Caco-2 cell line is APC mutated [29]. These mutations result in constitutively active Wnt signaling, supporting the high level of TOPFlash activation. When transfecting NCI-H508 cells, that carry the *VTI1A-TCF7L2* gene fusion (Fig 1a) [13], with the TOPFlash reporter plasmid, the luciferase activity was not observed to be significantly higher than that of the FOPFlash reporter plasmid (Fig 1b). These results suggest that either the NCI-H508 cell line does not have active Wnt signaling, or the VTI1A-TCF4 fusion protein in the NCI-H508 cells is not able to activate transcription from the TCF binding sites.

To determine whether the absence of TOPFlash activation in NCI-H508 cell was a result of the properties of the VTI1A-TCF4 fusion protein, LS174T cells were co-transfected with a plasmid expressing the VTI1A-TCF4 fusion protein and the TOPFlash plasmid. A 3-fold decrease in luciferase activity was observed when co-transfecting with VTI1A-TCF4 fusion protein expression plasmid ($p<0.0001$, Fig 1c). This shows that the presence of the VTI1A-TCF4 fusion protein in the LS174T cells, which has two functional *TCF7L2* alleles, has the ability to significantly reduce the strong Wnt signaling in LS174T cells, and indicates that the VTI1A-TCF4 fusion protein has dominant negative properties.

The *VTI1A* promoter is highly active compared to the *TCF7L2* promoter in LS174T cells

In order to investigate promoter activity of *VTI1A* and *TCF7L2*, reporter plasmid constructs containing the *VTI1A* promoter (pGL4.10-VTI1A-735) and the *TCF7L2* promoter (pGL4.10-TCF7L2-1405) were generated. These two constructs (Fig 2a) were transfected into the LS174T cell line. The luciferase activity of pGL4.10-VTI1A-735 was significantly higher, ~170-fold, compared with the activity of the pGL4.10 empty vector ($p<0.0001$) showing that the promoter is transcriptionally active in LS174T cells (Fig 2b). The activity of pGL4.10-TCF7L2-1405 was likewise significantly higher than pGL4.10 with a >40-fold increase in luciferase activity ($p<0.0001$). The luciferase activity of the construct containing the *VTI1A* promoter, pGL4.10-VTI1A-735, was significantly higher ($p<0.0001$) than the *TCF7L2* promoter construct. This shows that the promoter region for the *VTI1A-TCF7L2* fusion gene, encoding the VTI1A-TCF4 fusion protein, is much more transcriptionally active than the promoter for

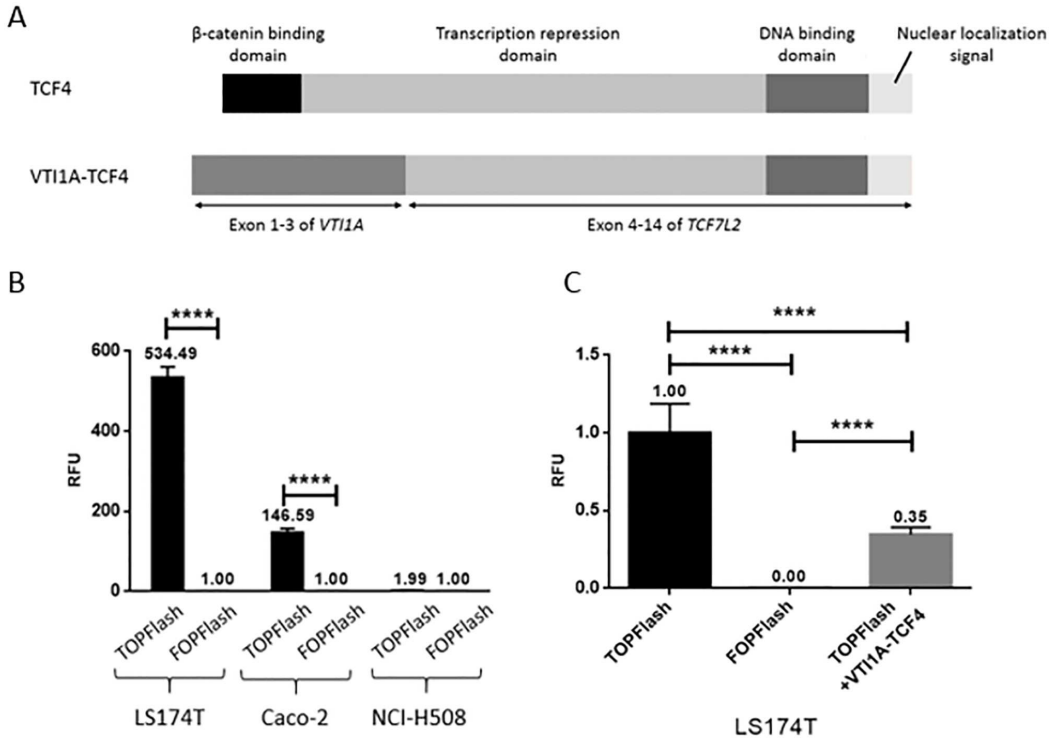


Fig 1. Wnt signaling in colon cancer cell lines. a Wildtype TCF4 contains a β-catenin binding domain, a transcription repression domain, a DNA binding domain, and a nuclear localization signal. With the fusion with VTI1A the resulting fusion protein, VTI1A-TCF4, contains the first three exons of VTI1A and lacks the β-catenin binding domain of TCF4. The rest of the TCF4 domains are still present. **b** The cells lines LS174T, Caco-2 and NCI-H508 were transfected with the TOPFlash and FOPFlash reporter plasmids, and luciferase activity was measured. Activity is stated in mean values relative to FOPFlash activity for each cell line and corrected with β-galactosidase activity. Error bars indicate SD, **** indicates $p<0.0001$, n = 4 **c** The LS174T cell line, showing high Wnt activity, was transfected with the TOPFlash and co-transfected with a plasmid expressing the VTI1A-TCF4 fusion protein. Activity is stated in mean values relative to TOPFlash activity and corrected with β-galactosidase activity. Error bars indicate SD, **** indicates $p<0.0001$, n = 8.

https://doi.org/10.1371/journal.pone.0200215.g001

TCF7L2 in the LS174T cell line. This will most likely result in altered expression of VTI1A-TCF4 compared to TCF7L2 *in vivo*.

The transcriptional activity of the *VTI1A* promoter is regulated by CDX2 in cultured LS174T cells

Genome-wide CDX2 ChIP-seq data from LS174T cells [30], shows a CDX2 ChIP-seq peak in the *VTI1A* promoter region, and the JASPAR CORE database identifies two potential CDX2 binding sites within this peak [31]. To investigate the relevance of CDX2's role in the transcriptional regulation of the *VTI1A* promoter region, a CDX2 expression plasmid and the pGL4.10-VTI1A-735 reporter plasmid were co-transfected into LS174T cells. Overexpression of CDX2 results in a 2-fold increase ($p<0.0001$) in transcriptional activity of pGL4.10-VTI1A-735 (Fig 3a). These results indicate that CDX2 has an up-regulatory effect on the transcriptional activity of the *VTI1A* promoter. Further indication of CDX2 regulation on the transcriptional activity of the *VTI1A* promoter was established in a LS174T *CDX2* knockout cell line from Pinto et al.

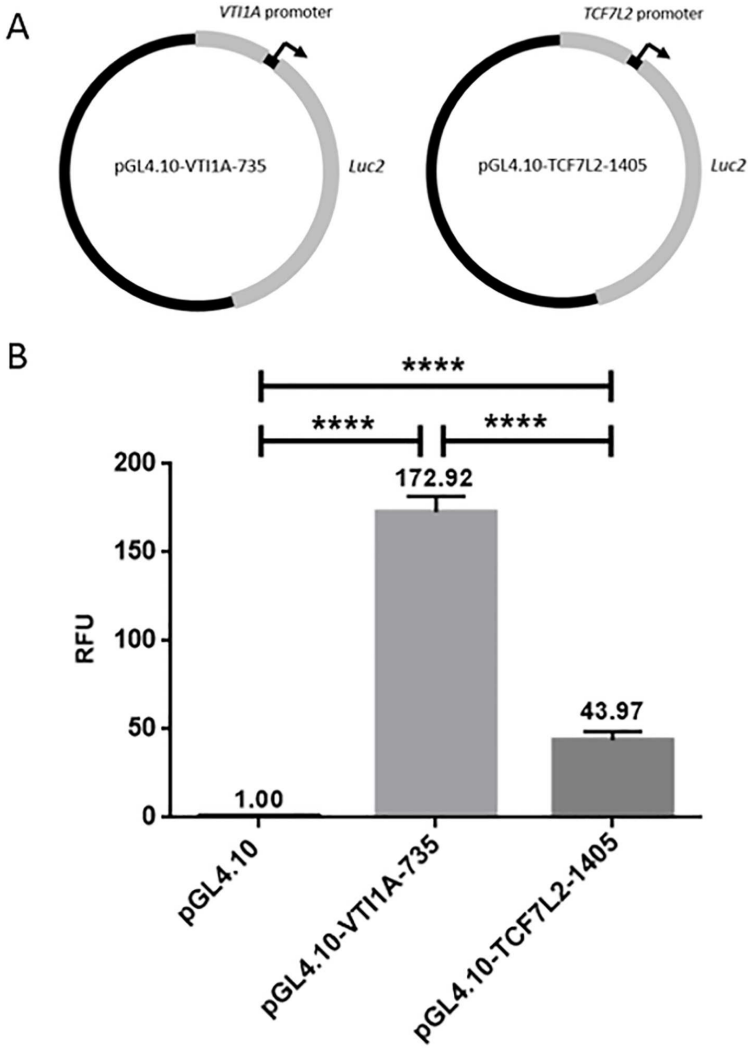


Fig 2. Transcriptional activity of the VTI1A and TCF7L2 promoters. **a** The pGL4.10-VTI1A-735 and pGL4.10-TCF7L2-1405 plasmids were created by inserting the promoter regions of VTI1A and TCF7L2, respectively, upstream the luciferase gene in the pGL4.10 plasmid. **b** LS174T cells were transfected with either the pGL4.10-VTI1A-375 or the pGL4.10-TCF7L2-1405 reporter plasmids. Activity is stated in mean values relative to pGL4.10 activity and corrected with β-galactosidase activity. Error bars indicate SD, **** indicates $p<0.0001$, $n = 8$.

https://doi.org/10.1371/journal.pone.0200215.g002

(2017). When transfecting the cells with the pGL4.10-VTI1A-735 reporter plasmid, the luciferase activity in the *CDX2* knockout cells decreased by 50% compared to the activity of the plasmid in the LS174T wildtype cells (Fig 3b).

To further elaborate on CDX2's role in the transcriptional regulation of the *VTI1A* promoter region, and thereby the expressional regulation of the VTI1A-TCF4 fusion protein,

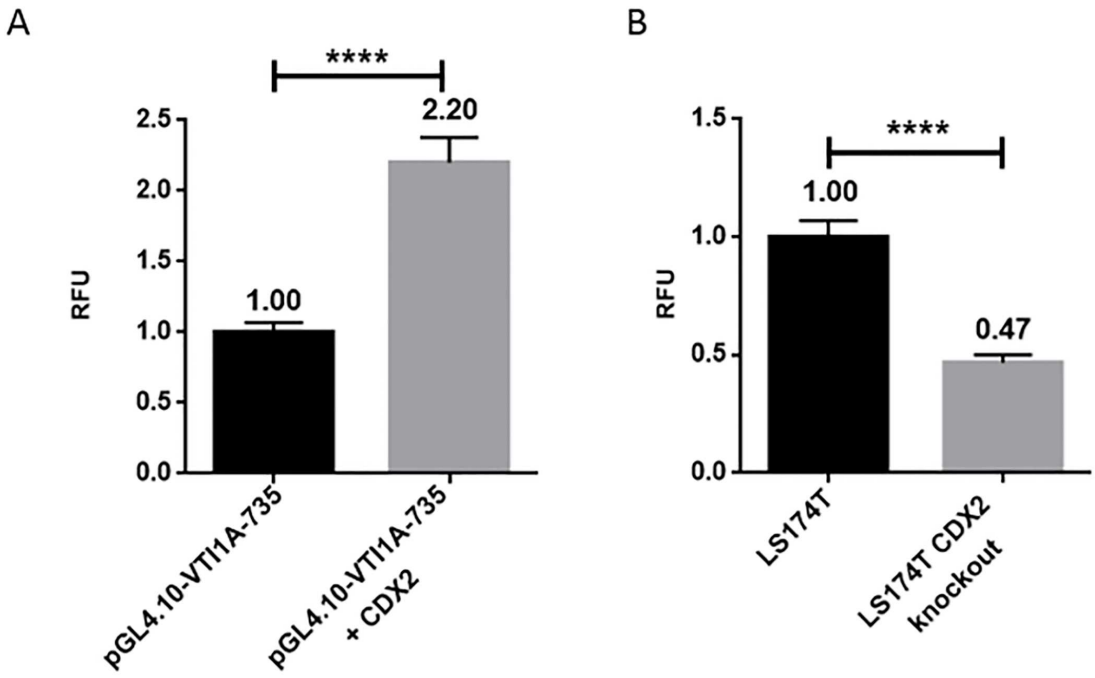


Fig 3. Transcriptional activity of CDX2. **a** LS174T cells were transfected with the pGL4.10-VTI1A plasmid-735 and co-transfected with a CDX2 expression plasmid. Activity is stated in mean values relative to pGL4.10 activity and corrected with β-galactosidase activity. Error bars indicate SD, **** indicates $p<0.0001$, n = 4. **b** The LS174T wild type cell line and a LS174T CDX2 knockout cell line was transfected with the pGL4.10-VTI1A plasmid. Activity is stated in mean values relative to pGL4.10 activity and corrected with β-galactosidase activity. Error bars indicate SD, **** indicates $p<0.0001$, n = 4.

https://doi.org/10.1371/journal.pone.0200215.g003

CDX2 ChIP assay on LS174T cells was performed, where chromatin-protein complexes were immunoprecipitated with CDX2-specific antibody or HA antibody, used as a negative control. The region investigated was determined from the CDX2 peak in the ChIP-seq data from Pinto et al. 2017, and can be seen on Fig 4a. The amount of CDX2 immunoprecipitated DNA was measured by qPCR using *VTI1A* promoter-specific primers flanking the potential binding region. The DNA was fragmented to a size of 300–400 bp in the ChIP assay. Thus, sequences flanking the CDX2 binding sites will also be precipitated. A good primer set for amplification of sequences nearby the ChIP-seq peak was chosen. Results show a significant enrichment of CDX2 ChIP DNA compared to the negative HA antibody control ($p<0.01$, Fig 4b). These results indicate that the CDX2 ChIP-seq peak in the *VTI1A* promoter contains CDX2 binding sites.

The validity of the identified CDX2 binding site as well as the two potential binding sites found using the JASPAR CORE database [31] was investigated through a deletion analysis using different lengths of the *VTI1A* promoter region cloned into the pGL4.10 plasmid (Fig 4c). The results of the deletion analysis show a 1.4-fold higher luciferase activity of the pGL4.10-VTI1A-532 plasmid compared to the activity of the pGL4.10-VTI1A-735 plasmid (Fig 4c). This increase in activity might be a result of inhibitory regions in the pGL4.10-VTI1A-735 plasmid. The luciferase activity of the pGL4.10-VTI1A-356 plasmid is 3.5-fold lower compared to the activity of the pGL4.10-VTI1A-532 plasmid, possibly an effect of this

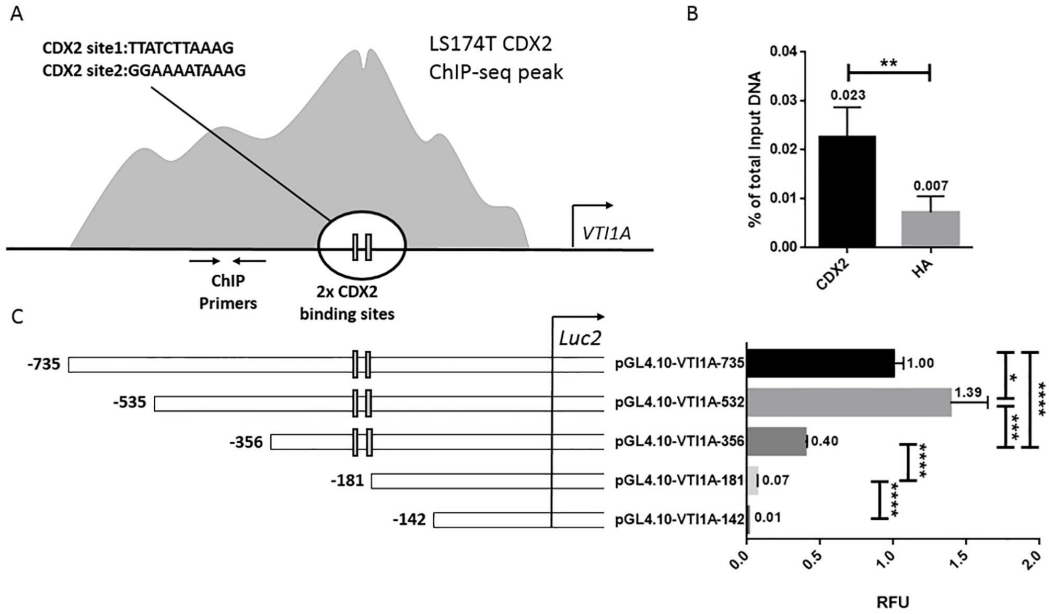


Fig 4. CDX2 regulation of *VTI1A* transcriptional activity. a LS174T CDX2 ChIP-seq data was used to identify a CDX2 peak in the VTI1A promoter region [30], and the JASPAR CORE database identifies two potential binding sites within this region [31]. **b** qPCR was performed on CDX2 immunoprecipitated DNA from LS174T cells with primers amplifying the 100 bp region seen in fig 4A. Error bars indicates SD, ** indicates $p<0.01$, n = 4 **c** Plasmids for deletion analysis with different lengths of the VTI1A promoter inserted in the pGl4.10 plasmid. The two possible CDX2 binding sites are marked. The deletion analysis was carried out in LS174T cells and the luciferase activity is stated in mean values relative to the pGL4.10-VTI1A-735 plasmid and is corrected with β-galactosidase activity. Error bars indicate SD, * indicates $p<0.05$, *** indicates $p<0.001$, **** indicates $p<0.0001$, n = 4.

https://doi.org/10.1371/journal.pone.0200215.g004

plasmid not containing the CDX2 regulatory region found previously. When comparing the luciferase activity of the pGL4.10-VTI1A-356 with that of the pGL4.10-VTI1A-181 plasmid a 5.7-fold decrease in activity and an almost complete abrogation of luciferase activity can be seen, indicating that fragment of *VTI1A* promoter omitted in the pGL4.10-VTI1A-181 plasmid is essential for the regulatory effect of CDX2 on the promoter. Furthermore, the CDX2 binding sites identified by the JASPER database are located in this fragment of the promoter region.

Discussion

Some members of the TCF/LEF family of transcription factors naturally occur as N-terminally truncated isoforms lacking the β-catenin binding site [22]. The nuclear localization signal, the DNA binding domain, and the transcription repression domain are still present in the truncated forms. This means that truncated isoforms of TCF/LEF are able to localize to the nucleus, bind to DNA binding sites, and recruit the Groucho co-repressor to the transcription repression domain. As the β-catenin binding site is not present in these truncated variants of TCF/LEF, β-catenin is unable to displace Groucho, resulting in repression of transcription [9,32]. Studies by Goncalves et al. and Van der Wetering et al. shows that constructs of truncated TCF4, that lack the N-terminal containing the β-catenin binding site, functions as dominant negative regulators of endogenous β-catenin/TCF complexes [23,24]. Our results show that the VTI1A-TCF4 fusion protein, which also lacks the N-terminal domain with β-catenin

binding domain, while it contains the nuclear localization signal, the DNA binding domain, and the transcription repression domain (Fig 1a), has dominant negative effects on WNT signaling.

Bass et al. (2011) did not investigate the regulatory properties of the VTI1A-TCF4 fusion protein, however they propose that it does not act as a full dominant negative protein as engineered dominant-negative *TCF4* alleles have been shown to strongly inhibit proliferation of colorectal carcinoma cell lines [24]. Furthermore, Bass et al. (2011) demonstrated that the VTI1A-TCF4 protein has a critical role in anchorage-independent growth of VTI1A-TCF4 fusion protein positive NCI-H508 cells, which would not have been expected if the VTI1A-TCF4 fusion protein is fully dominant negative [13]. In the study by Van der Wetering et al. (2002), doxycycline inducible plasmids expressing N-terminally truncated TCF4 were transfected into LS174T cells, and upon doxycycline induction the activity of TOPFlash was completely abrogated and the proliferation of the cells was inhibited [24]. However, when expressing the VTI1A-TCF4 fusion protein in LS174T cells, a complete abrogation of TOPFlash activity was not seen. This indicates that the fusion protein might exhibit different functional effects than that of a constructed truncated TCF4 protein and might therefore still allow for cell proliferation. Furthermore, the VTI1A-TCF4 fusion protein contains domains encoded by the *VTI1A* gene, which might add other functionalities to the VTI1A-TCF4 fusion protein.

Our results show that the *VTI1A* promoter region is significantly more active in LS174T cells than the promoter region for *TCF7L2*. This could mean increased expression of the VTI1A-TCF4 fusion protein in the proliferating cells of colonic crypts, where TCF4 is normally present [8]. The *VTI1A* promoter region was also shown to be regulated by CDX2 which is highly expressed in the upper two-thirds of the crypts [17,33]. Our results suggest that the VTI1A-TCF4 fusion protein will be expressed in the upper two-thirds of the crypts as well as at the base of the crypts. Studies have shown that TCF4 plays an important part in maintaining the proliferative cells in the colonic crypts, and that natural down-regulation of TCF4 expression as cells move up the colonic crypts may trigger colonic epithelial cell differentiation [6–8]. *Tcf7l* knockout mice die within 24 hours after birth and show no proliferating crypt stem cell compartments, supporting the idea that TCF4 plays an essential role in maintaining the proliferative compartment in the crypt [34].

The main binding partner of β-catenin in the colon is TCF4, but colon tumors isolated from *APC* mutated mice have an increased expression of LEF1 and TCF1 [3]. During active Wnt signaling, β-catenin will accumulate in the cell nucleus and with no binding site in the VTI1A-TCF4 fusion protein, β-catenin might bind and activate other members of the TCF/LEF family of transcription factors, such as TCF1 and LEF1. As the TCF/LEF family of transcription factors are indicated to work in distinct and opposing ways to regulate and maintain the equilibrium in cell proliferation and differentiation [1,9], a shift in this balance may contribute to the development or progression of colorectal cancer. A proposed model for the expression pattern of transcription factors in the mutated crypts can be seen in Fig 5.

In conclusion, we have determined that the VTI1A-TCF4 fusion protein transcribed from *VTI1A-TCF7L2* gene fusion has dominant negative properties and the transcription of this protein is activated by CDX2 in cultures LS174T cells. The role of the VTI1A-TCF4 fusion protein in colorectal cancer still needs to be investigated. However, as dysregulation of the Wnt signaling pathway plays an important role in the development of colorectal cancer the shift from wild type TCF4, that can function as a transcriptional activator, to the dominant negative VTI1A-TCF4 in mutated colonic cells will most likely have an effect on colon homeostasis and may perhaps contribute to the development of colorectal cancer.

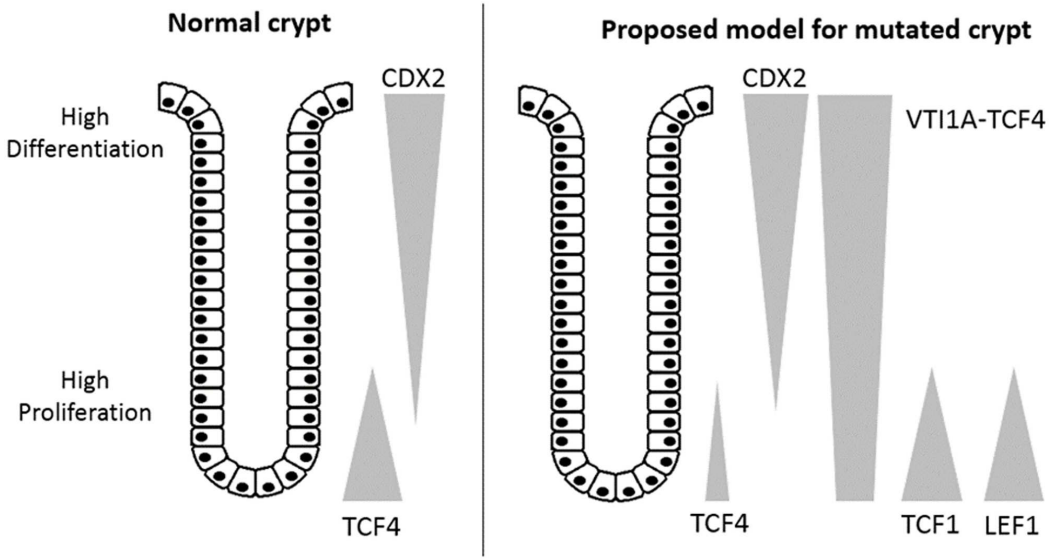


Fig 5. Expression in the colonic crypts. In normal crypts TCF4 is expressed at the base of the crypts [8] and as cells migrate towards the top of the crypts the expression of TCF4 decreases while expression of CDX2 increases [17,33]. In a proposed model for a crypt with cells carrying the *VTI1A-TCF7L2* fusion, the VTI1A-TCF4 fusion protein will be expressed throughout the crypt. As β-catenin is not able to bind to VTI1A-TCF4 it may instead bind and activate other members of the TCF/LEF family of transcription factors, e.g. TCF1 and LEF1.

https://doi.org/10.1371/journal.pone.0200215.g005

Supporting information

S1 Fig. Protein expression of VTI1A-TCF4. LS174T cells were transfected with the VTI1A-TCF4 expression plasmid and after 48 hours protein was extracted and used for Western Blot. The blot was incubated with TCF4 antibody. In cells transfected with the VTI1A-TCF4 expression plasmid a clear band can be seen at approx. 43 kDa, corresponding to the length of the VTI1A-TCF4 fusion protein. In the cells not transfected with the VTI1A-TCF4 expression plasmid, this band cannot be seen. For both samples the wild type TCF4 protein is seen at approx. 75 and 55 kDa.
(PDF)

S1 Table. Primers. List of primers used in the study.
(DOCX)

Author Contributions

Conceptualization: Johanne Davidsen, Katja Dahlgaard, Jesper Thorvald Troelsen.

Data curation: Johanne Davidsen, Jesper Thorvald Troelsen.

Formal analysis: Johanne Davidsen, Jesper Thorvald Troelsen.

Funding acquisition: Katja Dahlgaard, Jesper Thorvald Troelsen.

Investigation: Johanne Davidsen, Sylvester Larsen, Mehmet Coskun, Jesper Thorvald Troelsen.

Methodology: Johanne Davidsen, Eric Paul Bennett, Jesper Thorvald Troelsen.

Project administration: Jesper Thorvald Troelsen.

Resources: Jesper Thorvald Troelsen.

Supervision: Ismail Gögenur.

Validation: Johanne Davidsen.

Writing – original draft: Johanne Davidsen, Jesper Thorvald Troelsen.

Writing – review & editing: Johanne Davidsen, Sylvester Larsen, Mehmet Coskun, Ismail Gögenur, Katja Dahlgaard, Eric Paul Bennett, Jesper Thorvald Troelsen.

References

1. Kriegl L, Horst D, Reiche JA, Engel J, Kirchner T, Jung A. LEF-1 and TCF4 expression correlate inversely with survival in colorectal cancer. J Transl Med. 2010; 8: 1–8.
2. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-tieulent J, Jemal A. Global Cancer Statistics, 2012. Cancer J. 2015; 65: 87–108. https://doi.org/10.3322/caac.21262 PMID: 25651787
3. Angus-Hill ML, Elbert KM, Hidalgo J, Capecchi MR. T-cell factor 4 functions as a tumor suppressor whose disruption modulates colon cell proliferation and tumorigenesis. Proc Natl Acad Sci U S A. 2011; 108: 4914–4919. https://doi.org/10.1073/pnas.1102300108 PMID: 21383188
4. Olsen J, Espersen MLM, Jess P, Kirkeby L, Troelsen JT. The clinical perspectives of CDX2 expression in colorectal cancer: A qualitative systematic review. Surg Oncol. 2014; 23: 167–176. https://doi.org/10.1016/j.suronc.2014.07.003 PMID: 25126956
5. Fevr T, Robine S, Louvard D, Huelsken J. Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. Mol Cell Biol. 2007; 27: 7551–7559. https://doi.org/10.1128/MCB.01034-07 PMID: 17785439
6. Pinto D, Gregorieff A, Begthel H, Clevers H. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. Genes Dev. 2003; 17: 1709–1713. https://doi.org/10.1101/gad.267103 PMID: 12865297
7. Kuhnert F, Davis CR, Wang H, Chu P, Lee M, Yuan J, et al. Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. Proc Natl Acad Sci U S A. 2004; 101: 266–271. https://doi.org/10.1073/pnas.2536800100 PMID: 14695885
8. Mariadason JM, Bordonaro M, Aslam F, Shi L, Kuraguchi M, Velcich A, et al. Down-regulation of beta-catenin TCF signaling is linked to colonic epithelial cell differentiation. Cancer Res. 2001; 61: 3465–3471. PMID: 11309309
9. Hovanes K, Li TW, Munguia JE, Truong T, Milovanovic T, Lawrence Marsh J, et al. Beta-catenin-sensitive isoforms of lymphoid enhancer factor-1 are selectively expressed in colon cancer. Nat Genet. 2001; 28: 53–57. PMID: 11326276
10. Erdfelder F, Hertweck M, Filipovich A, Uhrmacher S, Kreuzer KA. High lymphoid enhancer- binding factor-1 expression is associated with disease progression and poor prognosis in chronic lymphocytic leukemia. Hematol Rep. 2010; 2: 2–5. https://doi.org/10.4081/hr.2010.e3 PMID: 22184516
11. Li Y, Wang L, Zhang M, Melamed J, Liu X, Reiter R, et al. LEF1 in androgen-independent prostate cancer: regulation of androgen receptor expression, prostate cancer growth, and invasion. Cancer Res. 2009; 69: 3332–3338. https://doi.org/10.1158/0008-5472.CAN-08-3380 PMID: 19351848
12. Roose J, Huls G, van Beest M, Moerer P, van der Horn K, Goldschmeding R, et al. Synergy between tumor suppressor APC and the beta-catenin-Tcf4 target Tcf1. Science (80-). 1999; 285: 1923–1926. https://doi.org/10.1126/science.285.5435.1923
13. Bass AJ, Lawrence MS, Brace LE, Ramos AH, Drier Y, Cibulskis K, et al. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. Nat Genet. 2011; 43: 964–968. https://doi.org/10.1038/ng.936 PMID: 21892161
14. Nome T, Hoff AM, Bakken AC, Rognum TO, Nesbakken A, Skotheim RI. High frequency of fusion transcripts involving TCF7L2 in colorectal cancer: Novel fusion partner and splice variants. PLoS One. Public Library of Science; 2014; 9: e91264. https://doi.org/10.1371/journal.pone.0091264 PMID: 24608966
15. Boyd M, Hansen M, Jensen TGK, Perearnau A, Olsen AK, Bram LL, et al. Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). J Biol Chem. 2010; 285: 25115–25125. https://doi.org/10.1074/jbc.M109.089516 PMID: 20551321

16. Verzi M., Hatzis P, Sulahian R, Philips J, Schuijers J, Shin H, et al. TCF4 and CDX2, major transcription factors for intestinal function, converge on the same cis-regulatory regions. Proc Natl Acad Sci U S A. 2010; 107: 15157–15162. https://doi.org/10.1073/pnas.1003822107 PMID: 20696899

17. Stegmann A, Hansen M, Wang Y, Larsen JB, Lund LR, Ritié L, et al. Metabolome, transcriptome, and bioinformatic cis-element analyses point to HNF-4 as a central regulator of gene expression during enterocyte differentiation. Physiol Genomics. 2006; 27: 141–155. https://doi.org/10.1152/physiolgenomics.00314.2005 PMID: 16868071

18. Olsen AK, Boyd M, Danielsen ET, Troelsen JT. Current and emerging approaches to define intestinal epithelium-specific transcriptional networks. AJP Gastrointest Liver Physiol. 2012; 302: 277–286. https://doi.org/10.1152/ajpgi.00362.2011 PMID: 22094602

19. Hryniuk A, Grainger S, Savory JGA, Lohnes D. Cdx1 and Cdx2 Function as Tumor Suppressors. J Biol Chem. 2014; 289: 33343–33354. https://doi.org/10.1074/jbc.M114.583823 PMID: 25320087

20. Natoli M, Christensen J, El-Gebali S, Felsani A, Anderle P. The role of CDX2 in Caco-2 cell differentiation. Eur J Pharm Biopharm. 2013; 85: 20–25. https://doi.org/10.1016/j.ejpb.2013.03.018 PMID: 23958315

21. Salari K, Spulak ME, Cuff J, Forster AD, Giacomini CP, Huang S. CDX2 is an amplified lineage-survival oncogene in colorectal cancer. PNAS Plus. 2012; 109: 3196–3205. www.pnas.org/cgi/doi/10.1073/pnas.1206004109

22. Van de Wetering M, Castrop J, Korinek V, Clevers H. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. Mol Cell Biol. 1996; 16: 745–752. PMID: 8622675

23. Gonçalves V, Matos P, Jordan P. The beta-catenin/TCF4 pathway modifies alternative splicing through modulation of SRp20 expression. RNA. 2008; 14: 2538–2549. https://doi.org/10.1261/rna.1253408 PMID: 18952824

24. Van de Wetering M, Sancho E, Verweij C, De Lau W, Oving I, Hurlstone A, et al. The B-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. Cell. 2002; 111: 241–250. https://doi.org/10.1016/S0092-8674(02)01014-0 PMID: 12408868

25. Veeman MT, Slusarski DC, Kaykas A, Louie SH, Moon RT. Zebrafish Prickle, a Modulator of Noncanonical Wnt/Fz Signaling, Regulates Gastulation Movements. Curr Biol. 2003; 13: 680–685.

26. Coskun M, Olsen AK, Holm TL, Kvist PH, Nielsen OH, Riis LB, et al. TNF-α-induced down-regulation of CDX2 suppresses MEP1A expression in colitis. Mol Basis Dis. 2012; 1822: 843–851. https://doi.org/10.1016/j.bbadis.2012.01.012 PMID: 22326557

27. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25: 402–408. https://doi.org/10.1006/meth.2001.1262 PMID: 11846609

28. Li VS., Ng SS, Boersema PJ, Low TY, Karthaus WR, Gerlach JP, et al. Wnt Signaling through Inhibition of β-Catenin Degradation in an Intact Axin1 Complex. Cell. 2012; 149: 1245–1256. https://doi.org/10.1016/j.cell.2012.05.002 PMID: 22682247

29. Hidalgo IJ, Raub TJ, Borchardt RT. Characterization of the human colon carcinoma cell line (Caco-2) as a model system for intestinal epithelial permeability. Gastroenterology. 1989; 96: 736–749. PMID: 2914637

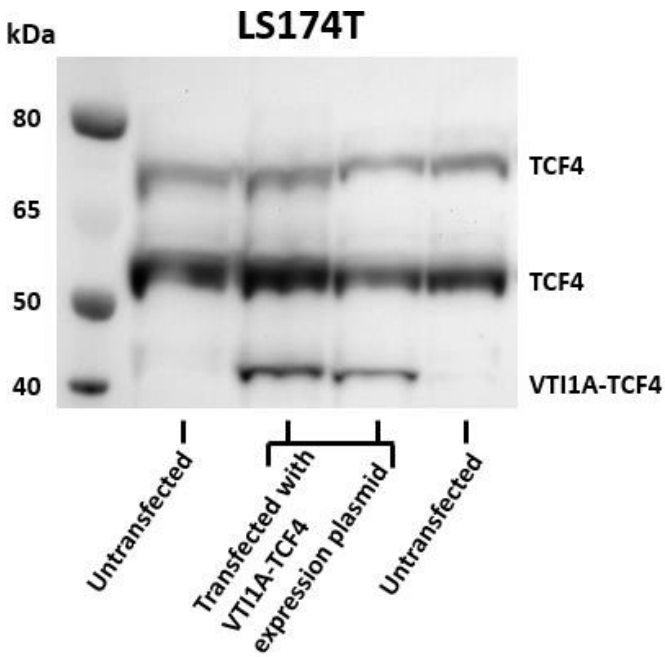
30. Pinto R, Hansen L, Hintze J, Almeida R, Larsen S, Coskun M, et al. Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression. Nucleic Acids Res. 2017; 45: 1–15.

31. Mathelier A, Fornes O, Arenillas DJ, Chen CY, Denay G, Lee J, et al. JASPAR 2016: A major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2016; 44: 110–115. https://doi.org/10.1093/nar/gkv1176 PMID: 26531826

32. Li TWH, Ting JT, Yokoyama NN, Bernstein A, van de Wetering M, Waterman ML. Wnt activation and alternative promoter repression of LEF1 in colon cancer. Mol Cell Biol. 2006; 26: 5284–5299. https://doi.org/10.1128/MCB.00105-06 PMID: 16809766

33. Olsen AK, Coskun M, Bzorek M, Kristensen MH, Danielsen ET, Jørgensen S, et al. Regulation of APC and AXIN2 expression by intestinal tumor suppressor CDX2 in colon cancer cells. Carcinogenesis. 2013; 34: 1361–1369. https://doi.org/10.1093/carcin/bgt037 PMID: 23393221

34. Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ, et al. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat Genet. 1998; 19: 379–383. https://doi.org/10.1038/1270 PMID: 9697701



S1 Fig. Protein expression of VTI1A-TCF4. LS174T cells were transfected with the VTI1ATCF4 expression plasmid and after 48 hours protein was extracted and used for Western Blot. The blot was incubated with TCF4 antibody. In cells transfected with the VTI1A-TCF4 expression plasmid a clear band can be seen at approx. 43 kDa, corresponding to the length of the VTI1A-TCF4 fusion protein. In the cells not transfected with the VTI1A-TCF4 expression plasmid, this band cannot be seen. For both samples the wild type TCF4 protein is seen at approx. 75 and 55 kDa.

S1 Table. Primers. List of primers used in the study

Primer set	Forward	Reverse
Promoter cloning		
VTI1A-735	GCCAGGTTGGTCTCAACATT	AAGCCTAGGTCAGGGAAAGG
VTI1A-532	CAGCCCAGGACAAGCAGAGAGC	AAGCCTAGGTCAGGGAAAGG
VTI1A-356	CAAGCTGCTGTGCCTTGTAG	AAGCCTAGGTCAGGGAAAGG
VTI1A-181	GGAAAATAAAGCACGCACGC	AAGCCTAGGTCAGGGAAAGG
VTI1A-142	GAGATTGCGACGAACAACCA	AAGCCTAGGTCAGGGAAAGG
TCF7L2-1405	TCGGGGATTTCATTTCAG	TTGGGGGTCTTTTTCTCC
ChIP		
CDX2 ChIP	TTCGTGTTCTGGGAATCGGC	TCTGTTACTGTCACTTCCGGA
cDNA synthesis		
VTI1A-TCF4	ACCATAAACAAAGCAGCGGG	GAGATTGCGACGAACAACCA

Paper III

CDX2 expression and perioperative patient serum affects the adhesion properties of cultured colon cancer cells

Davidsen J, Jessen SB, Watt SK, Larsen S, Dahlgaard K, Kirkegaard T, Gögenur I, and Troelsen JT.

BMC Cancer 2020, 20(426)

Surgery of patients with colon cancer induces a stress response that may be related to a higher risk of disease recurrence. In this study, we used a real-time, cell-based method to examine the effects of patient serum from before and 24 hours after colon cancer surgery on the adhesion abilities of cultured colon cancer cells. We also investigated the role of CDX2 on the adhesion abilities of colon cancer cells both pre- and postoperatively.

This study demonstrates a significant increase in adhesion abilities of cells in postoperative serum compared to preoperative serum, and further, that this increased adhesion postoperatively was dependent on CDX2 expression. These results may aid in determining the role prognostic abilities of CDX2 expression, as well as its role in metastasis development.

Davidsen et al. BMC Cancer (2020) 20:426
https://doi.org/10.1186/s12885-020-06941-y

BMC Cancer

RESEARCH ARTICLE Open Access

CDX2 expression and perioperative patient serum affects the adhesion properties of cultured colon cancer cells

Johanne Davidsen[1,2], Stine Bull Jessen[1,2], Sara Kehlet Watt[2], Sylvester Larsen[1,3], Katja Dahlgaard[1], Tove Kirkegaard[2], Ismail Gögenur[2] and Jesper T. Troelsen[1*]

Abstract

Background: Colon cancer is one of the most commonly diagnosed types of cancer with surgical resection of the tumor being the primary choice of treatment. However, the surgical stress response induced during treatment may be related to a higher risk of recurrence. The aim of this study was to examine the effect of surgery on adhesion of cultured colon cancer cells with or without expression of the tumour suppressor CDX2.

Method: We enrolled 30 patients undergoing elective, curatively intended laparoscopic surgery for colon cancer in this study. Blood samples were drawn 1 day prior to surgery and 24 h after surgery. The samples of pre- and postoperative serum was applied to wild type colon cancer LS174T cells and CDX2 inducible LS174T cells and adhesion was measured with Real-Time Cell-Analysis iCELLigence using electrical impedance as a readout to monitor changes in the cellular adhesion.

Results: Adhesion abilities of wild type LS174T cells seeded in postoperative serum was significantly increased compared to cells seeded in preoperative serum. When seeding the CDX2 inducible LS174T cells without CDX2 expression in pre- and postoperative serum, no significant difference in adhesion was found. However, when inducing CDX2 expression in these cells, the adhesion abilities in pre- and postoperative serum resembled those of the LS174T wild type cell line.

Conclusions: We found that the adhesion of colon cancer cells was significantly increased in postoperative versus preoperative serum, and that CDX2 expression affected the adhesive ability of cancer cells. The results of this study may help to elucidate the pro-metastatic mechanisms in the perioperative phase and the role of CDX2 in colon cancer metastasis.

Keywords: Colon cancer, CDX2, Tumour suppressor, Surgical stress, Metastasis, Cell adhesion, Perioperative phase

* Correspondence: troelsen@ruc.dk
[1]Department of Science and Environment, Enhanced Perioperative Oncology (EPeOnc) Consortium, Roskilde University, Universitetsvej 1, 4000 Roskilde, Denmark
Full list of author information is available at the end of the article

© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Colon cancer is the third most commonly diagnosed cancer and the second most leading cause of cancer-related death, accounting for approximately 1 in 10 cancer cases and deaths [1]. Surgical resection of the tumour is the primary choice of treatment but, despite medical and surgical advances, the risk of recurrence in colonic cancer is up to 30% after curative resection [2]. Manipulation of the tumour during surgery results in an increase in the number of circulating tumour cells [3], and the operation can lead to a surgical stress response (SSR) resulting in reduced anti-tumoral defence [4], as well as an increase in factors favouring an oncogenic environment [5]. Excessive stimulation of cytokine production during the SSR is associated with the risk of postoperative metastasis [6], and pro-inflammatory cytokines, such as IL-1 and TNF-α have been shown to stimulate adhesion in circulating cancer cells [2].

Alternations in cell adhesion is believed to be critical in cancer metastasis [5]. For tumour cells to disengage from the primary tumour, adhesion is downregulated through modification of the cadherin-catenin complex [7, 8]. Further, the integrity of tight junctions that maintain cell polarity in normal epithelia is diminished by downregulation of claudins [9]. Attachment of metastatic cancer cells to distant tissues is mediated through expression of selectins [10], integrins [11] and members of the immunoglobulin superfamily [12]. Surgical trauma provokes an inflammatory reaction which results in release of cytokines that are shown to increase adhesion of colon carcinoma cells to metastatic sites [13, 14]. This effect of cytokines on the cell adhesion molecules expressed on cancer cells may contribute to the development of metastasis [15].

The transcription factor Caudal type Homeobox 2 (CDX2) is crucial for the homeostasis of the colonic epithelium [16], and has been shown to be at tumour suppressor [17–20]. Lack of CDX2 expression in colon cancer cells is associated with aggressive clinical behaviour and can be used as an adverse prognostic biomarker [21–25]. CDX2 has been reported to be downregulated in colon cancer cells in the invasive front of the tumour [16, 26]. The downregulation of CDX2, and thereby loss of intestinal identity, has been suggested to be a precursor for metastatic colon cancer to perform epithelial-to-mesenchyme transition (EMT) [16, 26, 27]. As the circulating colon cancer cells establish metastasis they undergo mesenchyme-to-epithelia transition (MET) and CDX2 expression is re-established, allowing for it to be used as a marker to determine the primary tumours colonic origin [26]. Overexpression of CDX2 in colon cancer cell lines has shown decreased mobility and dissemination of cancer cells, further implicating fluctuation of CDX2 expression in the metastatic process [28]. Alterations in CDX2 expression is based on mechanisms such as inflammation and epigenetic regulation, rather than mutations [21, 29].

Through a cell-based assay, measuring cancer cell adhesion in a colon cancer cell culture treated with serum obtained from patients before and after colon cancer surgery, we aimed to examine whether laparoscopic colon cancer surgery affects the adhesion of cancer cells and if CDX2 influences the adhesion abilities of cultured colon cancer cells.

Methods

Participants

From January to July 2016, consecutive patients undergoing elective, curatively intended laparoscopic surgery for colon cancer, stage I-III according to Union for International Cancer Control (UICC), at Zealand University Hospital were enrolled in this study. Patients receiving neoadjuvant radio- or chemotherapy, with known immune defects, or previous cancer history, were excluded. All eligible patients received information regarding purpose and methods of the study and were included after giving oral and written consent. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by The Danish National Committee on Health Research Ethics, Region Zealand (file no: 2008-58-0020) and approved by the Danish Data Protection agency (protocol: SJ567).

Setting

During the perioperative period, patients followed the standard of care for colon cancer in a setting of Enhanced Recovery After Surgery (ERAS), which has been described in detail for the department elsewhere [30]. There were no restrictions on pain management, and all patients were encouraged to take their regular medication after surgery. The choice of anesthetics was determined at a pre-anesthesia interview, and patients received universal anesthesia with either Total Intravenous Anesthesia or volatile inhalational. For induction of anesthesia, propofol 2–3 mg/kg and remifentanil or sufentanil were administered. Hereafter, all patients received a single intravenous dose of 240 mg gentamycin and 1 g metronidazole. Patients assigned to Total Intravenous Anesthesia received a continuous infusion of propofol supplemented remifentanil 0.5 μg/kg/min. Patients assigned to volatile inhalation received sevoflurane to a minimum alveolar concentration of 0.7–1.2 and remifentanil or repeated boli of sufentanil. Prior to extubation, ondansetrone 4 mg, sufentanil 0.4–0.6 μg/kg, and 1 g of paracetamol was given. Ropivacaine, 20 mL, was administered locally in the wounds.

Data collection and processing

Demographic data was collected through the electronic patient charts including age, gender, smoking status, body mass index (BMI), American Society of Anesthesiologist (ASA) scores, and Charlson Comorbidity Index. The UICC stage was based on pre-operative CT scans and histology results. Blood samples were taken the day prior to surgery, and approximately 24 h after surgery. Samples were collected in serum separation gel-tubes and left undisturbed at room temperature for 30 min to allow clotting. Hereafter, samples were centrifuged at 2330 g at 4° C for 10 min to remove the clot. The resulting supernatant was immediately transferred into Eppendorf tubes and kept at -80 °C until analysis.

Cell culture

The wild type human colon cancer cell lines Caco-2, DLD-1, SW480, LoVo, LS174T and a CDX2 inducible LS174T cell line were used in this study. The CDX2 inducible LS174T cell line is genetically modified and are CDX2 knockout but contain inducible elements that enable activation of CDX2 expression by addition of doxycycline to the growth media [31]. LS174T cell lines were obtained from Assoc. Prof. Eric Paul Bennett. All cell lines were cultured in Dulbecco's Modified Eagle's Medium (DMEM) with Ultraglutamine with 4.5 g/L Glucose (Lonza, Basel, Switzerland) supplemented with 10% Fetal Bovine Serum (HyClone by Fisher Scientific, Waltham, MA, USA) and Penicillin (100 units/mL) Streptomycin (100 μg/mL) (Gibco by Life Technologies, Carlsbad, CA, USA) The cell cultures were incubated at 37 °C in 5% CO2 and passaged every 3–4 days. The LS174T cells with inducible CDX2 were cultured in media with or without 4 ng/ml doxycycline to induce CDX2 expression.

Adhesion measurement

Real-Time Cell-Analysis (RTCA) iCELLigence (ACEA Biosciences, San Diego, CA, USA) was used to measure cell adhesion. The RTCA iCELLigence instrument uses electrical impedance as a readout to monitor changes in the cellular phenotype. The cell culture plates used in the instrument have electrodes placed at the bottom of each well, and cells attaching to the electrodes will lead to an increase in electrical impedance. The relative change in the electrical impedance is recorded as a dimensionless value termed Cell Index. The RTCA iCELLigence was set up using E-Plate L8 PET (ACEA Biosciences, San Diego, CA, USA) and cells in DMEM containing either 7% pre- or postoperative serum were added to each well in quadruplicates. For the LS174T cell line, $2*10^4$ cells were seeded in each well. For the Caco-2 cell line, $5*10^3$ cells were seeded in each well, while for the DLD-1, SW480, and LoVo, $1*10^4$ cells were seeded in each well. For LS174T cell with inducible CDX2, $2*10^4$ cells with or without 4 ng/ml doxycycline induced CDX2 expression were seeded in replicates in the E-plate L8 PET. The impedance was measured every 5 min and the difference in Cell Index at 60 min between cells seeded in preoperative and postoperative serum was calculated.

Western blot

Cells for protein extraction were seeded in 6-well plates at $5*10^5$ cells/well. After 24 h media was changed and LS174T cells with inducible CDX2 were added media with or without doxycycline. Cells were lysed after 72 h of doxycycline treatment by rising with cold PBS and incubated 5 min with 150 μl/well 1x RIPA lysis buffer (1x PBS, 300 mM NaCl, 1% Tergitol NP-40, 0.1% SDS, 0.5% 7-Deoxycholic acid sodium salt, 0.5 μM EDTA pH 8.0) with freshly added 1 mM DTT and 2 μl/ml protease inhibitor mix p8340 (Sigma-Aldrich, St. Louis, MO, USA). Lysate was centrifuged for 15 min at 12.000 g and 4 °C. Supernatant was stored at − 20 °C. Protein concentration was determined by Bradford analysis (Bio-Rad, Hercules, CA, USA).

For the analysis, 10 μg protein was mixed 1:4 (v/v) with Bolt loading buffer and 1:10 (v/v) with Bolt sample reducing agent (Thermo Fisher Scientific, Waltham, MA, USA). Samples were incubated at 95 °C for 5 min and loaded on a Bolt 4–12% Bis-Tris Plus gel (Thermo Fisher Scientific, Waltham, MA, USA) PageRuler prestained protein ladder was used as marker (Thermo Fisher Scientific, Waltham, MA, USA). SDS-PAGE was performed in 1X Bolt MOPS running buffer (Thermo Fisher Scientific, Waltham, MA, USA) for 30 min at 25 V, then 60 min at 120 V. The gel was transferred by wet-electrotransfer to PVDF membrane for 60 min at 25 V in 1X NuPage transfer buffer (Thermo Fisher Scientific, Waltham, MA, USA). The membrane was blocked with dry skim milk diluted to 5% in Wash buffer (1X TBS with 0.1% Tween-20) for 1 h at room temperature. The membrane was washed with Wash buffer 5 times for 3 min and incubated overnight at 4 °C with primary antibody diluted in 2.5% skim milk in Wash buffer. The membrane was then washed 5 times for 3 min and incubated with diluted secondary antibody for 1 h at room temperature. Before visualization, the membrane was washed 5 times for 3 min and then visualized by incubating with the ECL solution SuperSignal West Dura Extended Duration Substrate for 5 min (Thermo Fisher Scientific, Waltham, MA, USA). Antibodies used: CDX2 1:1000 (BioGenex, Freemont, CA, USA, MU392A-UC); Vinculin 1:5000 (Abcam, Cambridge, UK, ab129002); Goat anti-rabbit HRP 1:10,000 (Thermo Fisher Scientific, Waltham, MA, USA, 32260); Goat anti-mouse HRP 1:10,000 (Thermo Fisher Scientific, Waltham, MA, USA, 32230).

Statistical analysis

The paired Wilcoxon signed-rank test was used to determine statistical differences in the adhesion of cells with pre- and postoperative serum and the level of statistical significance was set at p-values < 0.01. The RTCA iCELLigence data analysis software 1.0 and Graphpad Prism 8 software were used for statistical analysis.

Results

In total, 38 patients were enrolled in the study. Seven patients were excluded due to post-operative complications and one patient was excluded due to benign disease. A total of 30 patients, 19 male and 11 female, went through laparoscopic colon cancer surgery within an ERAS setting and were included in the study (see Table 1 for patient demographics). According to UICC staging [32], patients were diagnosed with stage I-III cancer. The patients had an ASA score [33] ranging from I to III and had between 0 and 2 in WHO Performance Status [34]. None of the patients had visible metastasis preoperatively. Serum from blood samples drawn on the day prior to surgery and the day after surgery was used for the analysis of adhesion.

Culturing five different colon cancer cell lines, LS174T, Caco-2, DLD-1, SW480, and LoVo, in media supplemented with perioperative serum from a single patient, showed increased adhesion abilities in cells seeded in postoperative serum compared to preoperative serum for all cell lines (Fig. 1a). The difference in Cell Index in percentage at 60 min varied from 3.5% in the LS174T cell line to 8.0% in the LoVo cell line (Fig. 1b). While all the cell lines showed varied extent of increase in adhesion in postoperative serum, we chose the LS174T cell line for testing our entire patient cohort consisting of 30 patients. This cell line was chosen as a genetically modified clone has been produced, which contains inducible elements that control the expression of CDX2 [31]. As a result, the cells do not express CDX2 without being induced. To our knowledge, this is the only colon cancer cell line still viable with complete depletion of CDX2 expression. In other CDX2 positive colon cancer cell lines, CDX2 acts as a linage survival gene that cannot be inactivated [35].

When investigating our cohort of 30 patients a significant difference in cell adhesion, with increased adhesion in wild type LS174T cells seeded in postoperative serum compared to preoperative serum was observed. A difference between the pre- and postoperative samples could be observed 20 min after seeding, and at 60 min the cells had adhered to the surface and no further increase in adhesion could be observed. The Cell Indexes at 60 min were for 26 out of 30 patients higher in the postoperative sample compared to the preoperative sample (p < 0.0001) (Fig. 2a). Cell Indexes were slightly lower for three patients in the postoperative serum (Fig. 2b). The sera from one patient gave the same Cell Index before and after surgery.

Table 1 Demographics for patients undergoing laparoscopic colonic resection for colon cancer

Age, mean (SD)		67,6 (8,8)
Gender n (%)	Male	19 (63,3)
	Female	11 (36,7)
ASA-score n (%)	1	3 (10,0)
	2	24 (80,0)
	3	3 (10,0)
BMI n (%)	< 18,5	1 (3,3)
	18.5–24.9	12 (40,0)
	25–29.9	8 (26,7)
	> 30	9 (30,0)
Smoking n (%)	Current smoker	5 (16,7)
	Former smoker	13 (43,3)
	Never smoker	12 (40,0)
Alcohol (drinks/week) n (%)	0–14/21	25 (83,3)
	> 14/21	5 (16,7)
Charlson Comorbidity Index n (%)	0	18 (60,0)
	1	6 (20,0)
	2	3 (10,0)
	Missing	3 (10,0)
WHO Performance status n (%)	0	25 (83,3)
	1	3 (10,0)
	2	2 (6,7)
UICC n (%)	1	10 (33,3)
	2	12 (40,0)
	3	8 (26,7)
Anesthesia n (%)	Intravenous	20 (66,7)
	Inhalation	10 (33,3)
Laparoscopic procedure n (%)	Right hemicolectomy	9 (30,0)
	Transverse colectomy	1 (3,3)
	Left hemicolectomy	1 (3,3)
	Sigmoidectomy	18 (60,0)
	Complete colectomy	1 (3,3)

ASA American Society of Anesthesiologist Score, *BMI* Body Mass Index, *UICC* Union for International Cancer Control

To investigate the role of CDX2 in cell adhesion, the colon cancer cell line LS174T with inducible CDX2 was used. This cell line has previously been used to study the effect of CDX2 on intestinal transcriptional regulation [36–38]. Western blotting analysis of the LS174T wild type and LS174T with inducible CDX2 cells was performed to detect CDX2 levels. Results show no CDX2 expression in the LS174T with inducible CDX2 when not treated with doxycycline (Fig. 3a). When treated

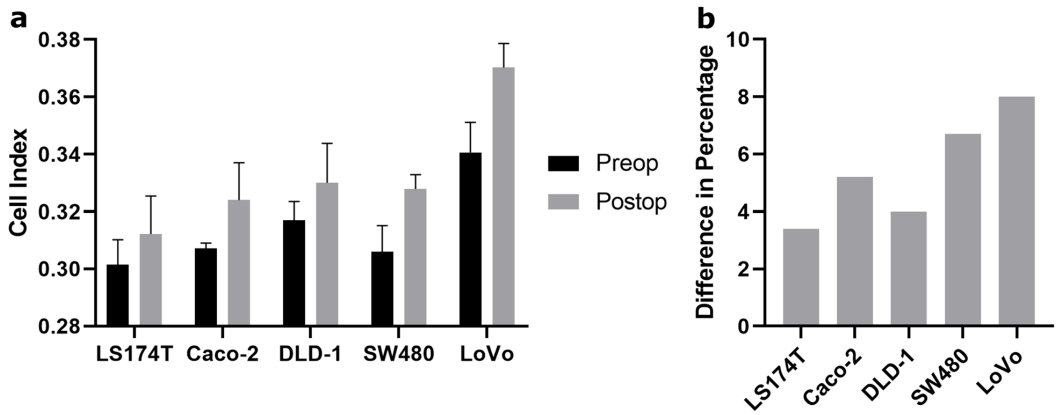


Fig. 1 Adhesion measurements of five different colon cancer cell lines in pre- or postoperative patient serum **a.** Cell adhesion of LS174T, Caco-2, DLD-1, SW480, and LoVo cells seeded in media with pre- or postoperative serum from one patient was measured. Mean Cell Index at 60 min is shown, $n = 4$. **b.** The difference in percentage between adhesion ability of cells seeded in postoperative serum compared to preoperative serum at 60 min was calculated for each cell line. The positive bars (grey) indicate higher adhesion in cells in postoperative serum compared to cells in preoperative serum

with doxycycline, expression of CDX2 was re-established. Vinculin was used as a control to measure total protein loaded. When seeding CDX2 negative LS174T cells in pre- and postoperative patient serum no difference in adhesion between the two groups was seen ($p = 0.21$) (Fig. 3b). Out of the 30 patient samples, only 11 had increased adhesion for cells seeded in postoperative serum compared to preoperative serum (Fig. 3c). However, when the cells had induced CDX2 at wild type levels, the results resembled those seen in the wild type LS174T cells, with significantly increased adhesion in postoperative serum compared to pre-operative serum ($p < 0.0001$) (Fig. 3d). Twenty-six patients out of 30 showed increased adhesion for cells seeded in postoperative serum compared to preoperative serum (Fig. 3e).

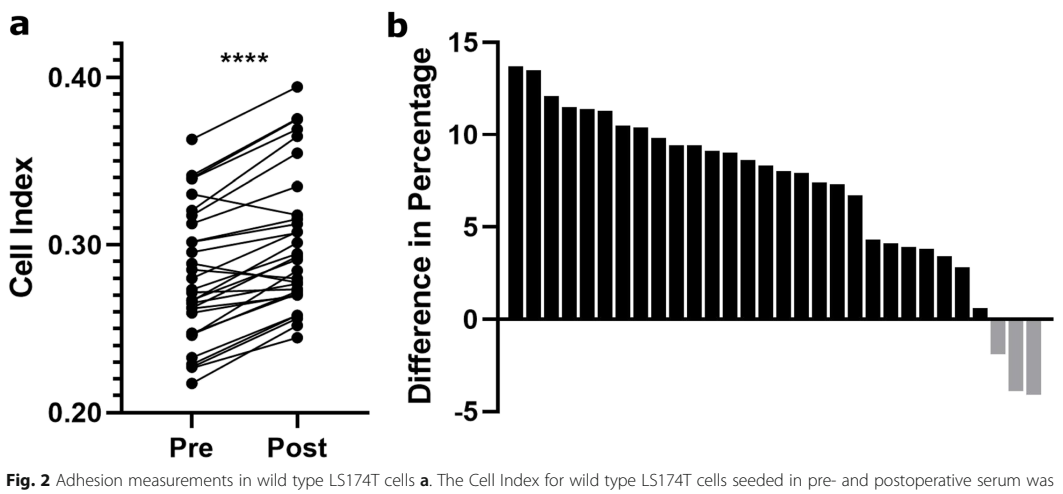


Fig. 2 Adhesion measurements in wild type LS174T cells **a**. The Cell Index for wild type LS174T cells seeded in pre- and postoperative serum was measured for each patient. Mean results at 60 min for pre- and postoperative cell adhesion for each patient is shown. ****$p < 0.0001$. **b.** The difference in percentage in adhesion at 60 min was calculated for each patient**.** The positive bars (black) indicate patients with higher adhesion in cells in postoperative compared to preoperative serum, while the negative bars (grey) indicate patients with higher adhesion in cells in preoperative compared to postoperative serum

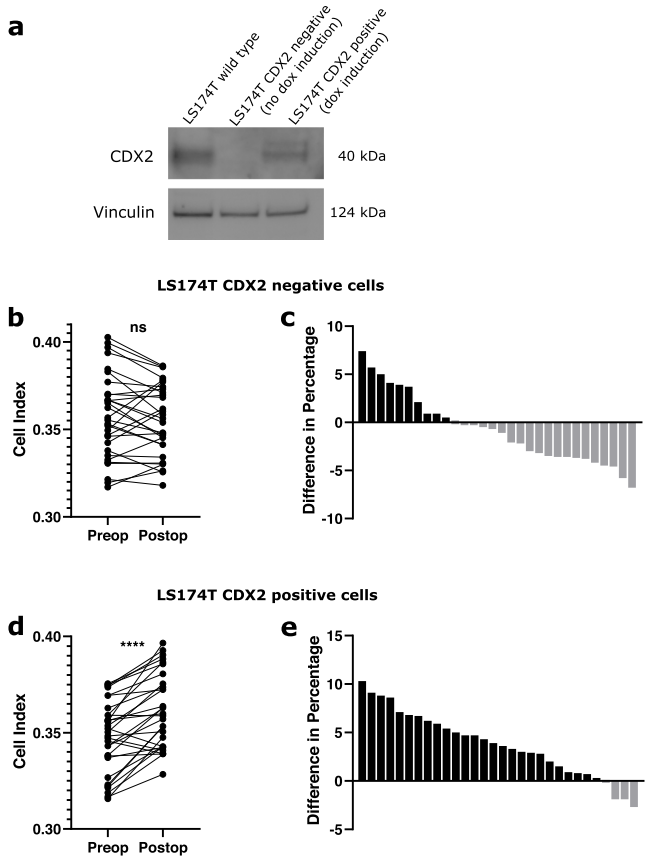


Fig. 3 Adhesion measurements in CDX2 inducible LS174T cells **a.** CDX2 protein expression was compared using western blotting. Cell lysate from LS174T wild type cells and LS174T cells with inducible CDX2 with or without doxycycline treatment was used in the analysis. Vinculin was used as a control. Bands are from the same gel **b.** The Cell Index for CDX2 negative LS174T cells seeded in pre- and postoperative serum was measured. Mean results at 60 min for pre- and postoperative cell adhesion for each patient is shown. n.s. = not significant. **c.** The difference in percentage in adhesion at 60 min was calculated for each patient**.** The positive bars (black) indicate patients with higher adhesion in postoperative compared to preoperative serum, while the negative bars (grey) indicate patients with higher adhesion in preoperative compared to postoperative serum. **d.** The Cell Index for CDX2 positive LS174T cells seeded in pre- and postoperative serum was measured. Mean results at 60 min for pre- and postoperative cell adhesion for each patient is shown. ****$p < 0.0001$. **e.** The difference in percentage in adhesion at 60 min was calculated for each patient**.** The positive bars (black) indicate patients with higher adhesion in postoperative compared to preoperative serum, while the negative bars (grey) indicate patients with higher adhesion in preoperative compared to postoperative serum

When comparing cell adhesion in cells treated with preoperative serum samples, there was a significant increase in adhesion in the CDX2 negative cells compared to the CDX2 positive cells ($p < 0.001$) (Fig. 4a). For the cells treated with postoperative serum, the CDX2 positive cells had significantly increased adhesion compared to the CDX2 negative cells ($p < 0.001$) (Fig. 4b).

Discussion

In this study, we established an in vitro method for measuring the effect of perioperative factors on the adhesion ability of the LS174T colon cancer cell line using serum from patients undergoing colon cancer surgery. Commonly used methods for cell adhesion assays typically include staining attached cells and using fluorescence for endpoint measurements [39–41], but by using the method developed in this paper, it is possible to monitor real-time cell adhesion for the entire adhesion period. While this method does not allow us to distinguish between initial sedimentation, cell attachment, cell spreading and stable cell adhesion, the mentioned are all part of the passive cell adhesion process [42].

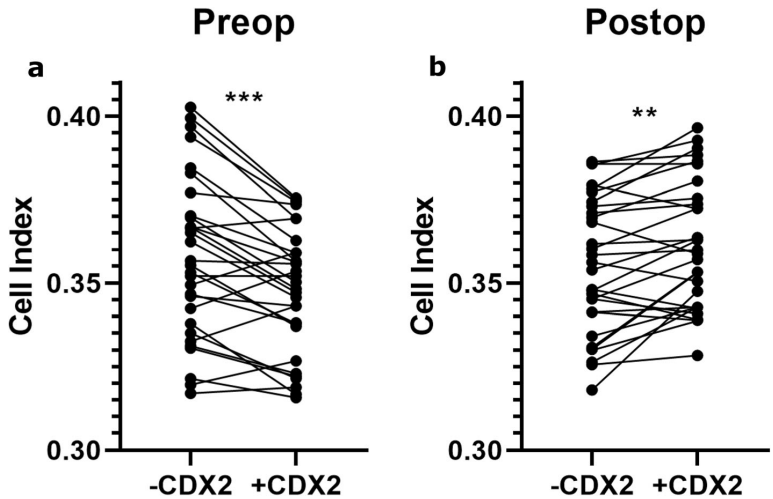


Fig. 4 Adhesion measurement of CDX2 inducible LS174T cells in pre- and postoperative serum **a.** The Cell Index for CDX2 negative and CDX2 positive inducible LS174T cells seeded in preoperative serum was measured. Mean results at 60 min for each patient is shown. ***$p < 0.001$. **b.** The Cell Index for CDX2 negative and CDX2 positive inducible LS174T cells seeded in postoperative serum was measured. Mean results at 60 min for each patient is shown. **$p < 0.01$

Investigating the adhesion abilities of the cells on a surface that more resembles the in vivo biological surface cancer cells interact with during metastasis may provide further insight to the adhesion process examined in this study.

Our study identified significantly increased cell adhesion abilities in five different colon cancer cell lines in postoperative serum, and further investigation using genetically modified LS174T cells showed this increase in adhesion to be eliminated by lack of CDX2 expression. This indicates that the absence of CDX2 expression results in reduced cancer cell adherence, and that fluctuation of CDX2 levels in cancer cells could be important in the metastatic process of colon cancer cells.

CDX2 has been shown to regulate the expression of a number of claudins [43, 44], a critical component of the tight junctions in epithelial cells. Aberrant expression of claudins has been seen in a variety of cancers, and it has been hypothesized that reduced claudin expression promotes tumorigenesis and metastasis by increasing the motility and invasion of cancer cells [9]. Studies have shown that reduced expression of claudin-1 is a predictor of poor prognosis and reduced disease-free survival [45–47], and that knockdown of claudin-1 expression in colon cancer cell lines significantly increase cell invasiveness [45]. Reduced expression of claudin-7 has been shown to be an early event in colorectal carcinogenesis [48], and downregulation of claudin-7 promotes EMT [49, 50]. Expression of claudin-23 has been shown to be downregulated in tumour tissue and downregulation is associated with shorter overall survival in patients with colorectal tumours [51]. Furthermore, CDX2 has been shown to mediate E-selectin ligand expression in colon cancer cells [52], a crucial component in the attachment of cancer cells to distant tissues during metastasis [10].

Exogenous CDX2 expression has been shown to be associated with reduced cell invasion in Lovo cells transfected with CDX2 overexpression plasmid [53], indicating that CDX2 may play a role in other metastatic processes besides adhesion. Other components have also been shown to influence the metastatic processes of adhesion, invasion and migration, such as the G-protein coupled receptor 55 [54], and the C-type lectin DC-SIGNR [55].

The observed difference in adhesion property between cells in pre- versus postoperative serum is most likely due to factors released into the bloodstream in patients during or after surgery. Previous studies have shown that pro-inflammatory cytokines mediate the adhesion of cancer cells to mesothelial and endothelial monolayers in vitro [13, 14]. Changes in expression of cell adhesion molecules in colon cancer cells have been associated with progression of cancer. This alteration in adhesion molecules could potentially facilitate the adhesion enabling intravasation as well as extravasation and may be part of organ selectivity in metastatic processes [15]. Furthermore, changes in expression of adhesion molecules could also affect postoperative cancer cell survival, as circulating tumour cells are vulnerable and depend on fast attachment in order to survive [2].

The underlying mechanisms of the interaction between the cellular adhesion molecules and factors in the patient serum has not yet been determined. However, when seeding the cells in patient serum, we can already measure altered adhesion abilities between cells in pre- vs postoperative serum 20 min after seeding. This rapid response indicates that factors in the patient serum directly affects the adhesion molecules already expressed on the surface of the cells or in the cytoplasm. Previously, RNA sequencing of the CDX2 inducible LS174T cell line used in our study showed altered RNA levels of several integrins, including integrin α3, α6, β4,and β6, in cells without CDX2 expression compared to wild-type LS174T cells [31]. Given the importance of the postoperative elevated adhesion and its possible correlation with recurrence, an identification of the precise mechanisms behind the interaction may provide valuable knowledge in reducing disease recurrence.

Conclusions

CDX2 expression is low in invasive colorectal cancer cells but is restored in metastases to a level corresponding to that of the primary tumour [26]. Our results show that CDX2 expression influences the adhesion ability of cultured colon cancer cells, and indicates that adjustments in CDX2 expression levels in cancer cells during EMT and MET is vital in the metastatic process of colon cancer. In conclusion, we demonstrate an in vitro method for measuring the effect of perioperative factors on the adhesion ability of the LS174T colon cancer cell line using serum form patients undergoing colon cancer surgery, and we demonstrate a differential effect on adhesion depending on CDX2 expression. If results from the method developed in this study can be shown to correlate with clinical oncological outcomes, the method may be applied in studies examining perioperative interventions in respect to their effect on short and long-term oncological outcomes after surgery.

Abbreviations

ASA: American society of Anaesthesiologists; BMI: Body Mass Index; CDX2: Caudal Type Homeobox 2; DMEM: Dulbecco's Modified Eagle's Medium; EMT: Epithelial-to-Mesenchyme Transition; ERAS: Enhanced Recovery After Surgery; MET: Mesenchyme-to-Epithelial Transition; RCTA: Real-Time Cell-Analysis; SSR: Surgical Stress Response; UICC: Union for International Cancer Control

Acknowledgements

We would like to thank Assoc. Professor Eric Paul Bennet for providing the cell lines used in this study.

Authors' contributions

JD, SBJ, SKW, SL, KD, TK, IG, and JTT conceived and designed the experiments. SKW, TK, and IG enrolled patients into the study and took blood samples. JD and SBJ performed the experiments. JD, SBJ, and JTT analyzed the data and performed statistical analysis. JD, SBJ, and SKW drafted the original manuscript. JD, SBJ, SKW, SL, KD, TK, IG, and JTT reviewed and edited the final manuscript. The author(s) read and approved the final manuscript.

Funding

The work has been supported by The Danish Council for Independent Research for Health and Disease, grant no. 4004-00140B. The funder had no role in the design of the study, in collection, analysis and interpretation of the data, or in writing the manuscript.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of The Danish National Committee on Health Research Ethics, Region Zealand (file no: 2008-58-0020), and approved by the Danish Data Protection agency (protocol: SJ567). Informed oral and written consent was obtained from all individual participants included in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Science and Environment, Enhanced Perioperative Oncology (EPeOnc) Consortium, Roskilde University, Universitetsvej 1, 4000 Roskilde, Denmark. [2]Center for Surgical Science, Enhanced Perioperative Oncology (EPeOnc) Consortium, Department of Surgery, Zealand University Hospital, Lykkebækvej 1, 4600 Køge, Denmark. [3]Department of Clinical Immunology, Naestved Hospital, Ringstedgade 77B, 4700 Naestved, Denmark.

Received: 17 January 2020 Accepted: 10 May 2020
Published online: 14 May 2020

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424 Available from: http://www.ncbi.nlm.nih.gov/pubmed/30207593.
2. van der Bij GJ, Oosterling SJ, Beelen RHJ, Meijer S, Coffey JC, van Egmond M. The perioperative period is an underutilized window of therapeutic opportunity in patients with colorectal cancer. Ann Surg. 2009;249(5):727–34.
3. Yamaguchi K, Takagi Y, Aoki S, Futamura M, Saji S. Significant detection of circulating cancer cells in the blood by reverse transcriptase-polymerase chain reaction during colorectal cancer resection. Ann Surg. 2000;232(1):58–65.
4. Neeman E, Ben-Eliyahu S. Surgery and stress promote cancer metastasis: new outlooks on perioperative mediating mechanisms and immune involvement. Brain Behav Immun. 2013;30:S32.
5. Läubli H, Borsig L. Altered cell adhesion and glycosylation promote Cancer immune suppression and metastasis. Front Immunol. 2019;10(September):1–9.
6. Horowitz M, Neeman E, Sharon E, Ben-eliyahu S. Exploiting the critical perioperative period to improve long-term cancer outcomes. Nat Publ Gr. 2015;12(April):213–26.
7. Shibata H, Takano H, Ito M, Shioya H, Hirota M, Matsumoto H, et al. Catenin is essential in intestinal adenoma formation. Proc Natl Acad Sci. 2007; 104(46):18199–204.
8. Aamodt R, Bondi J, Andersen SN, Bakka A, Bukholm G, Bukholm IRK. The prognostic impact of protein expression of E-cadherin-catenin complexes differs between rectal and Colon carcinoma. Gastroenterol Res Pract. 2010; 2010:1–7.
9. Tabariès S, Siegel PM. The role of claudins in cancer metastasis. Oncogene. 2017;36(9):1176–90.
10. Tremblay PL, Huot J, Auger FA. Mechanisms by which E-selectin regulates diapedesis of colon cancer cells under flow conditions. Cancer Res. 2008; 68(13):5167–76.
11. Robertson JH, Yang SY, Winslet MC, Seifalian AM. Functional blocking of specific integrins inhibit colonic cancer migration. Clin Exp Metastasis. 2009; 26(7):769–80.
12. Gallicchio M, Rosa AC, Dianzani C, Brucato L, Benetti E, Collino M, et al. Celecoxib decreases expression of the adhesion molecules ICAM-1 and

VCAM-1 in a colon cancer cell line (HT29). Br J Pharmacol. 2008;153(5):870–8.
13. Ten Kate M, Hofland LJ, van Grevenstein WMU, van Koetsveld PV, Jeekel J, van Eijck CHJ. Influence of proinflammatory cytokines on the adhesion of human colon carcinoma cells to lung microvascular endothelium. Int J Cancer. 2004 Dec;112(6):943–50.
14. Ziprin P, Ridgway PF, Pfister Uller KLM, Peck DH, Darzi AW. ICAM-1 mediated tumor-Mesothelial cell adhesion is modulated by IL-6 and TNF-α: a potential mechanism by which surgical trauma increases peritoneal metastases. Cell Commun Adhes. 2003;10:141–54.
15. Paschos KA, Canovas D, Bird NC. The role of cell adhesion molecules in the progression of colorectal cancer and the development of liver metastasis; 2009.
16. Olsen AK, Coskun M, Bzorek M, Kristensen MH, Danielsen ET, Jørgensen S, et al. Regulation of APC and AXIN2 expression by intestinal tumor suppressor CDX2 in colon cancer cells. Carcinogenesis. 2013 Jun;34(6):1361–9.
17. Hryniuk A, Grainger S, Savory JGA, Lohnes D. Cdx1 and Cdx2 function as tumor suppressors. J Biol Chem. 2014;289(48):33343–54.
18. Bonhomme C, Duluc I, Martin E, Chawengsaksophak K, Chenard M, Kedinger M, et al. The CDX2 homeobox gene has a tumour supressor function in the distal colon in addition to a homeotic role during gut development. Gut. 2003;52(10):1465–71.
19. Platet N, Hinkel I, Richert L, Murdamoothoo D, Moufok-Sadoun A, Vanier M, et al. The tumor suppressor CDX2 opposes pro-metastatic biomechanical modifications of colon cancer cells through organization of the actin cytoskeleton. Cancer Lett. 2017;386:57–64. Available from:. https://doi.org/10.1016/j.canlet.2016.10.040.
20. Balbinot C, Armant O, Elarouci N, Marisa L, Martin E, de Clara E, et al. The Cdx2 homeobox gene suppresses intestinal tumorigenesis through non-cell-autonomous mechanisms. J Exp Med. 2018;215(3):911–26.
21. Olsen J, Espersen MLM, Jess P, Kirkeby L, Troelsen JT. The clinical perspectives of CDX2 expression in colorectal cancer: a qualitative systematic review. Surg Oncol. 2014;23(3):167–76.
22. Bruun J, Sveen A, Barros R, Eide PW, Eilertsen I, Kolberg M, et al. Prognostic, predictive, and pharmacogenomic assessments of CDX2 refine stratification of colorectal cancer. Mol Oncol. 2018;12(9):1639–55.
23. Zhang BY, Jones JC, Briggler AM, Hubbard JM, Kipp BR, Sargent DJ, et al. Lack of caudal-type Homeobox transcription factor 2 expression as a prognostic biomarker in metastatic colorectal Cancer. Clin Colorectal Cancer. 2017;16(2):124–8. Available from:. https://doi.org/10.1016/j.clcc.2016.09.003.
24. Guo RJ, Funakoshi S, Lee HH, Kong J, Lynch JP. The intestine-specific transcription factor Cdx2 inhibits b-catenin/TCF transcriptional activity by disrupting the b-catenin-TCF protein complex. Carcinogenesis. 2010;31(2):159–66.
25. Dalerba P, Sahoo D, Paik S, Guo X, Yothers G, Song N, et al. CDX2 as a prognostic biomarker in stage II and stage III Colon Cancer. N Engl J Med. 2016;374(3):211–22. Available from:. https://doi.org/10.1056/NEJMoa1506597.
26. Brabletz T, Spaderna S, Kolb J, Hlubek F, Faller G, Bruns CJ, et al. Down-regulation of the Homeodomain factor Cdx2 in colorectal Cancer by collagen type I. Cancer Res. 2004;64(19):6973–7.
27. Zhang JF, Qu LS, Qian XF, Xia BL, Mao ZB, Chen WC. Nuclear transcription factor CDX2 inhibits gastric cancer-cell growth and reverses epithelial-to-mesenchymal transition in vitro and in vivo. Mol Med Rep. 2015;12(4):5231–8.
28. Gross I, Duluc I, Benameur T, Calon A, Martin E, Brabletz T, et al. The intestine-specific homeobox gene Cdx2 decreases mobility and antagonizes dissemination of colon cancer cells. Oncogene. 2008;27(1):107–15.
29. Subtil C, Guérin E, Schneider A, Chenard MP, Martin E, Domon-Dell C, et al. Frequent rearrangements and amplification of the CDX2 homeobox gene in human sporadic colorectal cancers with chromosomal instability. Cancer Lett. 2007;247(2):197–203.
30. Eriksen JR, Munk-Madsen P, Kehlet H, Gögenur I. Orthostatic intolerance in enhanced recovery laparoscopic colorectal resection. Acta Anaesthesiol Scand. 2019;63(2):171–7.
31. Pinto R, Hansen L, Hintze J, Almeida R, Larsen S, Coskun M, et al. Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression. Nucleic Acids Res. 2017;45:13.
32. Brierley JD, Gospodarowicz MK, Wittekind C. TNM Classification of Malignant Tumours. 8th ed. Hoboken: Wiley-Blackwell; 2016. p. 73–6.
33. Abouleish AE, Lieb ML, Cohen NH. ASA provides examples to each physical Status class. ASA Monit. 2015;79(6):28–49.
34. West HJ, Jin JO. Performance Status in Patients With Cancer. JAMA Oncol. 2015;1(7):998. Available from. https://doi.org/10.1001/jamaoncol.2015.3113.
35. Salari K, Spulak ME, Cuff J, Forster AD, Giacomini CP, Huang S. CDX2 is an amplified lineage-survival oncogene in colorectal cancer. PNAS Plus. 2012;109(46):3196–205.
36. Larsen S, Davidsen J, Dahlgaard K, Pedersen OB, Troelsen JT. HNF4α and CDX2 regulate intestinal YAP1 promoter activity. Int J Mol Sci. 2019 Jun;18:20(12).
37. Danielsen ET, Olsen AK, Coskun M, Nonboe AW, Larsen S, Dahlgaard K, et al. Intestinal regulation of suppression of tumorigenicity 14 (ST14) and serine peptidase inhibitor, Kunitz type −1 (SPINT1) by transcription factor CDX2. Sci Rep. 2018;8(1):1–14.
38. Davidsen J, Larsen S, Coskun M, Gögenur I, Dahlgaard K, Bennett EP, et al. The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2. PLoS One. 2018;13:7.
39. Sato Y, Kubo T, Morimoto K, Yanagihara K, Seyama T. High mannose-binding *Pseudomonas fluorescens* lectin (PFL) downregulates cell surface integrin/EGFR and induces autophagy in gastric cancer cells. BMC Cancer [Internet]. 2016;16(1):1–13. Available from: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L608079372%5Cnhttp://dx.doi.org/10.1186/s12885-016-2099-2%5Cnhttp://findit.library.jhu.edu/resolve?sid=EMBASE&issn=14712407&id=doi:10.1186/s12885-016-2099-2&atitle=High+mannose-bindi.
40. Ponda MP, Breslow JL. Serum stimulation of CCR7 chemotaxis due to coagulation factor XIIa-dependent production of high-molecular-weight kininogen domain 5. Proc Natl Acad Sci [Internet]. 2016;113(45):E7059–68. Available from:. https://doi.org/10.1073/pnas.1615671113.
41. Li J, Bowens N, Cheng L, Zhu X, Chen M, Hannenhalli S, et al. Myocardin-like protein 2 regulates TGF signaling in embryonic stem cells and the developing vasculature. Development. 2012;139(19):3531–42. Available from:. https://doi.org/10.1242/dev.082222.
42. Khalili AA, Ahmad MR. A review of cell adhesion studies for biomedical and biological applications. Int J Mol Sci. 2015;16(8):18149–84.
43. Sakaguchi T, Gu X, Golden HM, Suh E, Rhoads DB, Reinecker HC. Cloning of the human claudin-2 5′-flanking region revealed a TATA-less promoter with conserved binding sites in mouse and human for caudal-related homeodomain proteins and hepatocyte nuclear factor-1α. J Biol Chem. 2002;277(24):21361–70.
44. Satake S, Semba S, Matsuda Y, Usami Y, Chiba H, Sawada N, et al. Cdx2 transcription factor regulates claudin-3 and claudin-4 expression during intestinal differentiation of gastric carcinoma. Pathol Int. 2008;58(3):156–63.
45. Nakagawa S, Miyoshi N, Ishii H, Mimori K, Tanaka F, Sekimoto M, et al. Expression of CLDN1 in colorectal cancer: a novel marker for prognosis. Int J Oncol. 2011;39(4):791–6.
46. Shibutani M, Noda E, Maeda K, Nagahara H, Ohtani H, Hirakawa K. Low expression of claudin-1 and presence of poorly-differentiated tumor clusters correlate with poor prognosis in colorectal cancer. Anticancer Res. 2013;33(8):3301–6 Available from: http://www.ncbi.nlm.nih.gov/pubmed/23898096.
47. Yoshida T, Kinugasa T, Akagi Y, Kawahara A, Romeo K, Shiratsuchi I, et al. Decreased expression of claudin-1 in rectal cancer: a factor for recurrence and poor prognosis. Anticancer Res. 2011;31(7):2517–25 Available from: http://www.ncbi.nlm.nih.gov/pubmed/21873169.
48. Bornholdt J, Friis S, Godiksen S, Poulsen SS, Santoni-Rugiu E, Bisgaard HC, et al. The level of claudin-7 is reduced as an early event in colorectal carcinogenesis. BMC Cancer. 2011;11(1):65 Available from: http://www.biomedcentral.com/1471-2407/11/65.
49. Wang K, Li T, Xu C, Ding Y, Li W, Ding L. Claudin-7 downregulation induces metastasis and invasion in colorectal cancer via the promotion of epithelial-mesenchymal transition. Biochem Biophys Res Commun. 2019;508(3):797–804. Available from. https://doi.org/10.1016/j.bbrc.2018.10.049.
50. Bhat AA, Pope JL, Smith JJ, Ahmad R, Chen X, Washington MK, et al. Claudin-7 expression induces mesenchymal to epithelial transformation (MET) to inhibit colon tumorigenesis. Oncogene. 2015;34(35):4570–80 Available from: http://www.nature.com/articles/onc2014385.
51. Pitule P, Vycital O, Bruha J, Novak P, Hosek P, Treska V, et al. Differential expression and prognostic role of selected genes in colorectal cancer patients. Anticancer Res. 2013;33(11):4855–66.

52. Sakuma K, Aoki M, Kannagi R. Transcription factors c-Myc and CDX2 mediate E-selectin ligand expression in colon cancer cells undergoing EGF/bFGF-induced epithelial-mesenchymal transition. Proc Natl Acad Sci. 2012;109(20):7776–81.
53. Zheng JB, Sun XJ, Qi J, Li SS, Wang W, Ren HL, et al. Effects of homeodomain protein CDX2 expression on the proliferation and migration of Lovo colon cancer cells. Pathol Oncol Res. 2011;17(3):743–51.
54. Kargl J, Andersen L, Hasenöhrl C, Feuersinger D, Stančic A, Fauland A, et al. GPR55 promotes migration and adhesion of colon cancer cells indicating a role in metastasis. Br J Pharmacol. 2016;173(1):142–54.
55. Na H, Liu X, Li X, Zhang X, Wang Y, Wang Z, et al. Novel roles of DC-SIGNR in colon cancer cell adhesion, migration, invasion, and liver metastasis. J Hematol Oncol [Internet]. 2017;10(1):1–18. Available from:. https://doi.org/10.1186/s13045-016-0383-x.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC

5. Discussion

The CDX2 transcription factor is regarded to be the master regulator of intestinal identity of the colon and plays a vital part in maintaining intestinal homeostasis. It is responsible for the transcriptional regulation of many genes involved in normal colon function, including several genes involved in the Wnt signaling pathway, and dysregulation of CDX2 expression has been implicated in the development and progression of colon cancer. This thesis has investigated the role of CDX2 in intestinal homeostasis as well as its role in development and progression of colon cancer.

The development of a colon cancer cell line with inducible CDX2 expression described in **paper I** provides a model that can be used to investigate possible CDX2 target genes. Here we show that CDX2 regulates transcription of several novel genes, including genes involved in normal function of the colon, thereby expanding the role of CDX2 in colon homeostasis. A number of these new CDX2 targets are genes that have been associated with colon cancer, including the mucin encoded by the *MUC5AC*. The expression of *MUC5AC* in seen to be upregulated in early stage colon cancer, but decreases as the degree of dysplasia increases and patients with tumors that lack MUC5AC expression have worse prognosis (Imai et al., 2013; Wang et al., 2017). As *MUC5AC* is a CDX2 target gene, the dysregulation of CDX2 in colon cancer may be involved in both the upregulation of the MUC5CA mucin in the development of colon cancer, but also with the subsequent downregulation leading to a worse prognosis.

While other colon cancer cell lines are dependent on CDX2 expression (Natoli et al., 2013; Salari et al., 2012), the LS174T cell line with inducible CDX2 are seen to be viable despite complete lack of CDX2 expression. This indicates that CDX2 does not play as vital a part in this particular cell line as it does in others investigated. The mechanisms behind the viability of the cell line despite CDX2 knockout are unknown but the absence of CDX2 may lead to fundamental alternations in the signaling networks of the cells. Thus, this should be taken into consideration when interpreting results obtained using the LS174T colon cancer cell line with inducible CDX2.

In **paper II**, the VTI1A-TCF7L2 fusion protein was discovered to have dominant negative properties in Wnt signaling. This is not surprising, as other truncated members of the TCF/LEF family of transcription factors are seen to act as dominant negative regulators (Van de Wetering et al., 1996). However, we also uncover that the promoter of the fusion protein is transcriptionally regulated by CDX2, resulting in not only a protein with altered functional characteristics but also altered expression pattern. Other fusion proteins that play vital roles in cancer development have been identified, but only the functional characteristics of these proteins have been investigated (Soda et al., 2007; Tomlins et al., 2005; Yu et al., 2010). While altered protein function may be critical in oncogenesis, alternative promoter regions and thereby altered expression patterns may be an equally critical component in the effect of fusion proteins. As cancer cells are genetically instable, multiple fusion proteins have been observed that can potentially drive cancer development (Gao et al., 2018). When assessing the possible oncogenic effects of fusion proteins, not only the functional properties of the resulting fusion proteins should be investigated, but also the possible altered transcriptional regulation.

The functional characteristics of the VTI1A-TCF7L2 fusion protein was only investigated in relation to its effect on the Wnt signaling pathway. The effects of the truncated v-SNARE protein, encoded by *VTI1A,* were not investigated in **paper II.** The v-SNARE protein facilitates vesicle transport by fusion of intracellular vesicles from the late endosome to the trans-Golgi complex (Bass et al., 2011). Several genome-wide association studies have shown association between known variants of the *VTI1A* gene and colorectal cancer (Wang et al., 2014; Zeng et al., 2016), lung cancer (Su et al., 2015), and glioma (Kinnersley et al., 2015; Melin et al., 2017), and a fusion involving the *VTI1A* gene has also been discovered in hepatocellular carcinoma (Tsuge et al., 2019). Therefore, the truncated VTI1A protein in the VTI1A-TCF7L2 fusion protein may also play a role in colon cancer tumorigenesis and the effect of the altered v-SNARE should be further investigated.

In **paper III** we use a real-time, cell-based assay to show, for the first time, that perioperative patient serum affects the adhesion abilities of colon cancer cell lines. We

demonstrate that the adhesion abilities of cancer cells are increased in postoperative serum compared to preoperative serum, and that this effect on adhesion seems to be dependent on CDX2 expression. While the molecular mechanisms behind the role of CDX2 in adhesion are yet to be determined, CDX2 is known to regulate transcription of a number of molecules important in cell adhesion, including cadherins (Hinoi et al., 2002), claudins (Bhat et al., 2012; Sakaguchi et al., 2002; Satake et al., 2008), and matripase (Danielsen et al., 2018). Through RNA sequencing in **paper I** we have also discovered that the RNA levels of several integrins are altered in the inducible cells lacking CDX2 expression compared to wild-type cells, further implicating CDX2 in the cellular adhesion process.

The method used in **paper III** utilizes the ability of colon cancer cell lines to adhere to a tissue-culture treated plastic surface. While the results clearly show alternations in the cells ability to adhere, it would be beneficial to examine cell adhesion on a more biological surface. This may also further aid in determining the molecular mechanisms behind the role of CDX2 in cell adhesion.

The impact of the level of CDX2 on cellular adhesion reveals a possible mechanism by which CDX2 can be responsible for the progression of colon cancer. While the applicability of CDX2 as a prognostic marker is still debatable, the majority of studies find loss of CDX2 expression in cancer cells to be associated with worse overall survival (Bae et al., 2015; Bruun et al., 2018; Neumann et al., 2018). In tumor budding cells the level of CDX2 is often observed to be decreased (Coskun et al., 2014; Graule et al., 2018; Hansen et al., 2018), and when taking the results from **paper III,** showing that CDX2 affects adhesion abilities of colon cancer cells, into account, the downregulation of CDX2 may be necessary for the cancer cells to disengage from the tumor. It has previously been suggested that downregulation of CDX2 acts as a precursor to EMT, and as circulating cancer cells reach a metastatic site, the level of CDX2 expression increases and the cancer cells undergo MET (Brabletz et al., 2004; Zhang et al., 2015). Fluctuation in the level of CDX2 during the metastatic process may therefore be an important component in cancer metastasis.

DISCUSSION

CDX2 is broadly known as a tumor suppressor in colon cancer and loss of CDX2 expression in regarded to result in worse prognosis in colon cancer patients (Dalerba et al., 2016; Mallo et al., 1997). However, evidence also suggests that CDX2 is essential in cancer cell survival and proliferation (Natoli et al., 2013; Yu et al., 2019), thereby indicating that it may also have oncogenic properties. In other gastrointestinal cancers, as well as leukemia, CDX2 is regarded as an oncogene (Barros et al., 2012; Rawat et al., 2012; Tamagawa et al., 2012). This debate of CDX2 as either a tumor suppressor or an oncogene indicates that both downregulation of expression as well as overexpression may play a role in cancer development.

The results presented in this thesis have shown the substantial effect CDX2 has on the transcriptional activity of numerous genes in colon cancer cells. Further, it has demonstrated the influence of the expression of CDX2 on cell morphology in a perioperative setting, exploring its role in the metastatic progression of cancer.
Despite the many advances into uncovering the impact of CDX2 dysregulation in colon cancer development, its role is not yet fully understood. Neither is the effect of CDX2 expression on disease progression and metastasis, although studies suggest a possible prognostic ability. Thus, further investigations into the molecular mechanisms behind the role of CDX2 in the development and progression of colon cancer are important in order to improve future treatment of disease.

6. Concluding remarks

This thesis has investigated the effect of CDX2 expression on the transcriptional activity of genes involved in colon homeostasis and explored its role in the development and progression of colon cancer.

A number of novel CDX2 target genes important in the homeostasis of the colon were discovered in **paper I**, where a method to integrate inducible transcriptional elements into cultured cells was developed. This model allowed tight control of CDX2 expression, in turn enabling investigation into downstream effector genes under varying concentrations of CDX2.

The promoter of the recurrent colon cancer fusion protein VTI1A-TCF7L2 was identified in **paper II** as being transcriptionally activated by CDX2 in colon cancer cell lines. The fusion protein itself was shown to have dominant negative properties, resulting in a protein with not only altered functional characteristics but also altered expression pattern in the colon epithelium.

The effect of surgical stress and CDX2 expression on adhesion abilities of colon cancer cell lines was demonstrated using a real-time, cell-based method in **paper III**. Colon cancer cells showed increased adhesion abilities in postoperative serum compared to preoperative serum in patients undergoing colon cancer surgery. It was further shown that the increased adhesion in the postoperative serum was dependent on CDX2 expression.

In summary, the results of this thesis provide information into the role of CDX2 in colon homeostasis, as well as its implication in colon cancer development and progression. Several new target genes of CDX2 are uncovered, including the colon cancer fusion protein VTI1A-TCF7L2 that could play a role in cancer development. The adhesion abilities of cultured cancer cells are seen to be affected by CDX2 expression, associating dysregulation of CDX2 expression in the metastatic processes of colon cancer.

7. Future perspectives

While the results from the papers presented in the thesis show the possible implication of CDX2 in colon cancer progression, the molecular mechanism behind the effect of CDX2 in the metastatic process remains to be determined. The regulation of several cell-cell adhesion and tight junction molecules, including cadherins, claudins, and integrins, are seen to be regulated my CDX2 and modulation though the expression of these molecules may be the source of the effect on metastatic processes. Other cellular abilities, such as migration, invasion, and proliferation, are also considered important in the metastatic process, however the possible role of CDX2 in these are still unknown.

The studies presented in the three papers of this thesis all rely on colon cancer cell lines, in particular the LS174T cell line. This cell line, as well as the other cell lines used in the studies, are all well-established cell lines that have been utilized for many years to investigate different aspects of colon regulation and cancer development. However, the selection of cells capable of surviving in vitro has introduced genetic alterations and these alternations may make them fundamentally different from cancer cells in the human body. This may result in altered signaling pathways that can affect the results gained from using this type of model. Another type of cell-based model that can be used are three-dimensional in vitro colon organoids. They contain the different types of cells found in the colon and thereby represent a more physiological model of the colon. They can be isolated and grown from both healthy and tumor tissue samples, and allows for a more detailed overview of colon function. Despite the advantages, the organoids are technically more difficult to handle and the standardized assays used in 2D cell lines must be customized the growing conditions of organoids. However, with the development of new handling techniques and validation of assays to be used in organoid culture they may become the standard cell-based system in the future. Investigating the impact of CDX2 on colon cancer development and progression in this cell system would be interesting.

References

Aamodt, R., Bondi, J., Andersen, S.N., Bakka, A., Bukholm, G., Bukholm, I.R.K., 2010. The Prognostic Impact of Protein Expression of E-Cadherin-Catenin Complexes Differs between Rectal and Colon Carcinoma. Gastroenterol. Res. Pract. 2010, 1–7.

Ackerman, M.E., Chalouni, C., Schmidt, M.M., Raman, V. V., Ritter, G., Old, L.J., Mellman, I., Wittrup, K.D., 2008. A33 antigen displays persistent surface expression. Cancer Immunol. Immunother. 57, 1017–1027.

Angell, H.K., Bruni, D., Carl Barrett, J., Herbst, R., Galon, J., 2020. The immunoscore: Colon cancer and beyond. Clin. Cancer Res. 26, 332–339.

Angus-Hill, M.L., Elbert, K.M., Hidalgo, J., Capecchi, M.R., 2011. T-cell factor 4 functions as a tumor suppressor whose disruption modulates colon cell proliferation and tumorigenesis. Proc. Natl. Acad. Sci. U. S. A. 108, 4914–4919.

Asgari-Karchekani, S., Karimian, M., Mazoochi, T., Taheri, M.A., Khamehchian, T., 2019. CDX2 Protein Expression in Colorectal Cancer and ItsCorrelation with Clinical and Pathological Characteristics, Prognosis, and Survival Rate of Patients. J. Gastrointest. Cancer.

Baba, Y., Nosho, K., Shima, K., Freed, E., Irahara, N., Philips, J., Meyerhardt, J.A., Hornick, J.L., Shivdasani, R.A., Fuchs, C.S., Ogino, S., 2009. Relationship of CDX2 Loss with Molecular Features and Prognosis in Colorectal Cancer: Implications for Clinical and Pathology Practice. Clin Cancer Res. Clin Cancer Res. July 15, 4665–4673.

Bae, J.M., Lee, T.H., Cho, N.Y., Kim, T.Y., Kang, G.H., 2015. Loss of CDX2 expression is associated with poor prognosis in colorectal cancer patients. World J. Gastroenterol. 21, 1457–1467.

Balbinot, C., Armant, O., Elarouci, N., Marisa, L., Martin, E., de Clara, E., Onea, A., Deschamps, J., Beck, F., Freund, J.N., Duluc, I., 2018. The Cdx2 homeobox gene suppresses intestinal tumorigenesis through non-cell-autonomous mechanisms. J. Exp. Med. 215, 911–926.

Barros, R., Da Costa, L.T., Pinto-de-Sousa, J., Duluc, I., Freund, J.N., David, L., Almeida, R., 2011. CDX2 autoregulation in human intestinal metaplasia of the stomach: Impact on the stability of the phenotype. Gut 60, 290–298.

Barros, R., Freund, J.N., David, L., Almeida, R., 2012. Gastric intestinal metaplasia revisited: Function and regulation of CDX2. Trends Mol. Med. 18, 555–563.

Bass, A.J., Lawrence, M.S., Brace, L.E., Ramos, A.H., Drier, Y., Cibulskis, K., Sougnez, C., Voet, D., Saksena, G., Sivachenko, A., Jing, R., Parkin, M., Pugh, T., Verhaak, R.G., Stransky, N., Boutin, A.T., Barretina, J., Solit, D.B., Vakiani, E., Shao, W., Mishina, Y., Warmuth, M., Jimenez, J., Chiang, D.Y., Signoretti, S., Kaelin, W.G., Spardy, N., Hahn, W.C., Hoshida, Y., Ogino, S., DePinho, R.A., Chin, L., Garraway, L.A., Fuchs, C.S., Baselga, J., Tabernero, J., Gabriel, S., Lander, E.S., Getz, G., Meyerson, M., 2011. Genomic sequencing of colorectal adenocarcinomas

identifies a recurrent VTI1A-TCF7L2 fusion. Nat. Genet. 43, 964–968.

Benahmed, F., Gross, I., Gaunt, S.J., Beck, F., Jehan, F., Domon-Dell, C., Martin, E., Kedinger, M., Freund, J.N., Duluc, I., 2008. Multiple Regulatory Regions Control the Complex Expression Pattern of the Mouse Cdx2 Homeobox Gene. Gastroenterology 135, 1238–1247.

Bhat, A.A., Sharma, A., Pope, J., Krishnan, M., Washington, M.K., Singh, A.B., Dhawan, P., 2012. Caudal homeobox protein Cdx-2 cooperates with Wnt pathway to regulate claudin-1 expression in colon cancer cells. PLoS One 7.

Boman, B.M., Fields, J.Z., 2013. An APC:WNT Counter-Current-Like Mechanism Regulates Cell Division Along the Human Colonic Crypt Axis: A Mechanism That Explains How APC Mutations Induce Proliferative Abnormalities That Drive Colon Cancer Development. Front. Oncol. 3, 1–15.

Bonetti, L.R., Lionti, S., Vitarelli, E., Barresi, V., 2017. Prognostic significance of CDX2 immunoexpression in poorly differentiated clusters of colorectal carcinoma. Virchows Arch. 471, 731–741.

Boulanger, J., Vézina, A., Mongrain, S., Boudreau, F., Perreault, N., Auclair, B.A., Lainé, J., Asselin, C., Rivard, N., 2005. Cdk2-dependent phosphorylation of homeobox transcription factor CDX2 regulates its nuclear translocation and proteasome-mediated degradation in human intestinal epithelial cells. J. Biol. Chem. 280, 18095–18107.

Boyd, M., Bressendorff, S., Møller, J., Olsen, J., Troelsen, J.T., 2009. Mapping of HNF4alpha target genes in intestinal epithelial cells. BMC Gastroenterol. 9, 1–16.

Boyd, M., Hansen, M., Jensen, T.G.K., Perearnau, A., Olsen, A.K., Bram, L.L., Bak, M., Tommerup, N., Olsen, J., Troelsen, J.T., 2010. Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). J. Biol. Chem. 285, 25115–25125.

Brabletz, T., Spaderna, S., Kolb, J., Hlubek, F., Faller, G., Bruns, C.J., Jung, A., Nentwich, J., Duluc, I., Domon-Dell, C., Kirchner, T., Freund, J.-N., 2004. Down-Regulation of the Homeodomain Factor Cdx2 in Colorectal Cancer by Collagen Type I. Cancer Res. 64, 6973–6977.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA. Cancer J. Clin. 68, 394–424.

Brenner, H., Kloor, M., Pox, C.P., 2014. Colorectal cancer. Lancet 383, 1490–1502.

Bruun, J., Sveen, A., Barros, R., Eide, P.W., Eilertsen, I., Kolberg, M., Pellinen, T., David, L., Svindland, A., Kallioniemi, O., Guren, M.G., Nesbakken, A., Almeida, R., Lothe, R.A., 2018. Prognostic, predictive, and pharmacogenomic assessments of CDX2 refine stratification of colorectal cancer. Mol. Oncol. 12, 1639–1655.

Buzza, M.S., Netzel-Arnett, S., Shea-Donohue, T., Zhao, A., Lin, C.Y., List, K., Szabo, R., Fasano, A., Bugge, T.H., Antalis, T.M., 2010. Membrane-anchored serine protease matriptase regulates epithelial barrier formation and permeability in the intestine. Proc. Natl. Acad. Sci. U. S. A. 107, 4200–4205.

Cai, L., Bennedsen, A.L.B., Qvortrup, C., Gögenur, I., 2019. Increasing incidence of colorectal cancer in young patients. Ugeskr. Laeger 182.

REFERENCES

Cattin, A., Le Beyec, J., Barreau, F., Saint-Just, S., Houllier, A., Gonzalez, F.J., Robine, S., Pinçon-Raymond, M., Cardot, P., Lacasa, M., Ribeiro, A., 2009. Hepatocyte nuclear factor 4alpha, a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. Mol. Cell. Biol. 29, 6294–6308.

Cecchini, M.J., Walsh, J.C., Parfitt, J., Chakrabarti, S., Correa, R.J., MacKenzie, M.J., Driman, D.K., 2019. CDX2 and Muc2 immunohistochemistry as prognostic markers in stage II colon cancer. Hum. Pathol. 90, 70–79.

Cernat, L., Blaj, C., Jackstadt, R., Brandl, L., Engel, J., Hermeking, H., Jung, A., Kirchner, T., Horst, D., 2014. Colorectal cancers mimic structural organization of normal colonic crypts. PLoS One 9, 4–11.

Chan, D.S.M., Lau, R., Aune, D., Vieira, R., Greenwood, D.C., Kampman, E., Norat, T., 2011. Red and processed meat and colorectal cancer incidence: Meta-analysis of prospective studies. PLoS One 6.

Chawengsaksophak, K., James, R., Hammond, V.E., Köntgen, F., Beck, F., 1997. Homeosis and intestinal tumours in Cdx2 mutant mice. Nature 386, 84–87.

Coskun, M., Boyd, M., Olsen, J., Troelsen, J.T., 2010. Control of intestinal promoter activity of the cellular migratory regulator gene ELMO3 by CDX2 and SP1. J. Cell. Biochem. 109, 1118–1128.

Coskun, M., Olsen, A.K., Bzorek, M., Holck, S., Engel, U.H., Nielsen, O.H., Troelsen, J.T., 2014. Involvement of CDX2 in the cross talk between TNF-α and Wnt signaling pathway in the colon cancer cell line Caco-2. Carcinogenesis 35, 1185–1192.

Coskun, M., Olsen, A.K., Holm, T.L., Kvist, P.H., Nielsen, O.H., Riis, L.B., Olsen, J., Troelsen, J.T., 2012. TNF-α-induced down-regulation of CDX2 suppresses MEP1A expression in colitis. Mol. Basis Dis. 1822, 843–851.

Coskun, M., Soendergaard, C., Joergensen, S., Dahlgaard, K., Riis, L.B., Nielsen, O.H., Sandelin, A., Troelsen, J.T., 2017. Regulation of Laminin γ2 Expression by CDX2 in Colonic Epithelial Cells Is Impaired During Active Inflammation. J. Cell. Biochem. 118, 298–307.

Dalerba, P., Sahoo, D., Paik, S., Guo, X., Yothers, G., Song, N., Wilcox-Fogel, N., Forgó, E., Rajendran, P.S., Miranda, S.P., Hisamori, S., Hutchison, J., Kalisky, T., Qian, D., Wolmark, N., Fisher, G.A., van de Rijn, M., Clarke, M.F., 2016. CDX2 as a Prognostic Biomarker in Stage II and Stage III Colon Cancer. N. Engl. J. Med. 374, 211–222.

Danielsen, E.T., Olsen, A.K., Coskun, M., Nonboe, A.W., Larsen, S., Dahlgaard, K., Bennett, E.P., Mitchelmore, C., Vogel, L.K., Troelsen, J.T., 2018. Intestinal regulation of suppression of tumorigenicity 14 (ST14) and serine peptidase inhibitor, Kunitz type -1 (SPINT1) by transcription factor CDX2. Sci. Rep. 8, 1–14.

Davidsen, J., Jessen, S.B., Watt, S.K., Larsen, S., Dahlgaard, K., Kirkegaard, T., Gögenur, I., Troelsen, J.T., 2020. CDX2 expression and perioperative patient serum affects the adhesion properties of cultured colon cancer cells. BMC Cancer 20, 1–10.

REFERENCES

Davidsen, J., Larsen, S., Coskun, M., Gögenur, I., Dahlgaard, K., Bennett, E.P., Troelsen, J.T., 2018. The VTI1A-TCF4 colon cancer fusion protein is a dominant negative regulator of Wnt signaling and is transcriptionally regulated by intestinal homeodomain factor CDX2. PLoS One 13.

Dawson, H., Galván, J.A., Helbling, M., Muller, D.E., Karamitopoulou, E., Koelzer, V.H., Economou, M., Hammer, C., Lugli, A., Zlobec, I., 2014. Possible role of Cdx2 in the serrated pathway of colorectal cancer characterized by BRAF mutation, high-level CpG Island methylator phenotype and mismatch repair-deficiency. Int. J. Cancer 134, 2342–2351.

Dekker, E., Tanis, P.J., Vleugels, J.L.A., Kasi, P.M., Wallace, M.B., 2019. Colorectal cancer. Lancet 394, 1467–1480.

Dhawan, P., Singh, A.B., Deane, N.G., No, Y.R., Shiou, S.R., Schmidt, C., Neff, J., Washington, M.K., Beauchamp, R.D., 2005. Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer. J. Clin. Invest. 115, 1765–1776.

Drummond, F., Putt, W., Fox, M., Edwards, Y.H., 1997. Cloning and chromosome assignment of the human CDX2 gene. Ann. Hum. Genet. 61, 393–400.

Fearon, E.R., 2011. Molecular Genetics of Colorectal Cancer. Annu. Rev. Pathol. Mech. Dis. 6, 479–507.

Fedirko, V., Tramacere, I., Bagnardi, V., Rota, M., Scotti, L., Islami, F., Negri, E., Straif, K., Romieu, I., La Vecchia, C., Boffetta, P., Jenab, M., 2011. Alcohol drinking and colorectal cancer risk: An overall and dose-Response meta-analysis of published studies. Ann. Oncol. 22, 1958–1972.

Fevr, T., Robine, S., Louvard, D., Huelsken, J., 2007. Wnt/beta-catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. Mol. Cell. Biol. 27, 7551–7559.

Gallicchio, M., Rosa, A.C., Dianzani, C., Brucato, L., Benetti, E., Collino, M., Fantozzi, R., 2008. Celecoxib decreases expression of the adhesion molecules ICAM-1 and VCAM-1 in a colon cancer cell line (HT29). Br. J. Pharmacol. 153, 870–878.

Gao, N., White, P., Kaestner, K.H., 2009. Establishment of Intestinal Identity and Epithelial-Mesenchymal Signaling by Cdx2. Dev. Cell 16, 588–599.

Gao, Q., Liang, W.W., Foltz, S.M., Mutharasu, G., Jayasinghe, R.G., Cao, S., Liao, W.W., Reynolds, S.M., Wyczalkowski, M.A., Yao, L., et al., 2018. Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. Cell Rep. 23, 227-238.e3.

Garrison, W.D., Battle, M.A., Yang, C., Kaestner, K.H., Sladek, F.M., Duncan, S.A., 2006. Hepatocyte nuclear factor 4a is essential for embryonic development of the mouse colon. Gastroenterology 130, 1207–1220.

Gehart, H., Clevers, H., 2019. Tales from the crypt: new insights into intestinal stem cells. Nat. Rev. Gastroenterol. Hepatol. 16, 19–34.

Grady, W.M., Markowitz, S.D., 2016. The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening. Dig Dis Sci. 60, 762–772.

Graule, J., Uth, K., Fischer, E., Centeno, I., Galván, J.A., Eichmann, M., Rau, T.T., Langer, R., Dawson, H., Nitsche, U., Traeger, P., Berger, M.D., Schnüriger, B., Hädrich, M., Studer, P., Inderbitzin, D., Lugli, A., Tschan, M.P., Zlobec, I., 2018. CDX2 in colorectal cancer is an independent prognostic factor and regulated by promoter methylation and histone deacetylation in tumors of the serrated pathway. Clin. Epigenetics 10, 120.

Gross, I., Duluc, I., Benameur, T., Calon, A., Martin, E., Brabletz, T., Kedinger, M., Domon-Dell, C., Freund, J.N., 2008. The intestine-specific homeobox gene Cdx2 decreases mobility and antagonizes dissemination of colon cancer cells. Oncogene 27, 107–115.

Gross, I., Lhermitte, B., Domon-Dell, C., Duluc, I., Martin, E., Gaiddon, C., Kedinger, M., Freund, J.N., 2005. Phosphorylation of the homeotic tumor suppressor Cdx2 mediates its ubiquitin-dependent proteasome degradation. Oncogene 24, 7955–7963.

Guo, R.J., Funakoshi, S., Lee, H.H., Kong, J., Lynch, J.P., 2010. The intestine-specific transcription factor Cdx2 inhibits b-catenin/TCF transcriptional activity by disrupting the b-catenin-TCF protein complex. Carcinogenesis 31, 159–166.

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. Cell 144, 646–674.

Hansen, T.F., Kjær-Frifeldt, S., Eriksen, A.C., Lindebjerg, J., Jensen, L.H., Sørensen, F.B., Jakobsen, A., 2018. Prognostic impact of CDX2 in stage II colon cancer: results from two nationwide cohorts. Br. J. Cancer 119, 1367–1373.

Hao, H.X., Xie, Y., Zhang, Y., Zhang, O., Oster, E., Avello, M., Lei, H., Mickanin, C., Liu, D., Ruffner, H., Mao, X., Ma, Q., Zamponi, R., Bouwmeester, T., Finan, P.M., Kirschner, M.W., Porter, J.A., Serluca, F.C., Cong, F., 2012. ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. Nature 485, 195–202.

Heaney, A., Buggy, D.J., 2012. Can anaesthetic and analgesic techniques affect cancer recurrence or metastasis? Br. J. Anaesth. 109, i17–i28.

Herr, R., Köhler, M., Andrlová, H., Weinberg, F., Möller, Y., Halbach, S., Lutz, L., Mastroianni, J., Klose, M., Bittermann, N., Kowar, S., Zeiser, R., Olayioye, M.A., Lassmann, S., Busch, H., Boerries, M., Brummer, T., 2015. B-Raf inhibitors induce epithelial differentiation in BRAF-mutant colorectal cancer cells. Cancer Res. 75, 216–229.

Hinkel, I., Duluc, I., Martin, E., Guenot, D., Freund, J., Gross, I., 2012. Cdx2 controls expression of the protocadherin Mucdhl, an inhibitor of growth and B-catenin activity in colon cancer cells. Gastroenterology 142, 875-885.e3.

Hinoi, T., Gesina, G., Akyol, A., Kuick, R., Hanash, S., Giordano, T.J., Gruber, S.B., Fearon, E.R., 2005. CDX2-regulated expression of iron transport protein hephaestin in intestinal and colonic epithelium. Gastroenterology 128, 946–961.

Hinoi, T., Lucas, P.C., Kuick, R., Hanash, S., Cho, K.R., Fearon, E.R., 2002. CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia. Gastroenterology 123, 1565–1577.

Hinoi, T., Tani, M., Lucas, P.C., Caca, K., Dunn, R.L., Macri, E., Loda, M., Appelman,

H.., Cho, K.R., Fearon, E.R., 2001. Loss of CDX2 expression and microsatellite instability are prominent features of large cell minimally differentiated carcinomas of the colon. Am. J. Pathol. 159, 2239–2248.

Hoppler, S., Kavanagh, C.L., 2007. Wnt signalling: variety at the core. J. Cell Sci. 120, 385–393.

Houde, M., Laprise, P., Jean, D., Blais, M., Asselin, C., Rivard, N., 2001. Intestinal Epithelial Cell Differentiation Involves Activation of p38 Mitogen-activated Protein Kinase that Regulates the Homeobox Transcription Factor CDX2. J. Biol. Chem. 276, 21885–21894.

Hryniuk, A., Grainger, S., Savory, J.G.A., Lohnes, D., 2012. Cdx function is required for maintenance of intestinal identity in the adult. Dev. Biol. 363, 426–437.

Hryniuk, A., Grainger, S., Savory, J.G.A., Lohnes, D., 2014. Cdx1 and Cdx2 Function as Tumor Suppressors. J. Biol. Chem. 289, 33343–33354.

Imai, Y., Yamagishi, H., Fukuda, K., Ono, Y., Inoue, T., Ueda, Y., 2013. Differential mucin phenotypes and their significance in a variation of colorectal carcinoma. World J. Gastroenterol. 19, 3958–3968.

Jess, T., Rungoe, C., Peyrin-Biroulet, L., 2012. Risk of Colorectal Cancer in Patients With Ulcerative Colitis: A Meta-analysis of Population-Based Cohort Studies. Clin. Gastroenterol. Hepatol. 10, 639–645.

Jiang, Y., Ben, Q., Shen, H., Lu, W., Zhang, Y., Zhu, J., 2011. Diabetes mellitus and incidence and mortality of colorectal cancer: A systematic review and meta-analysis of cohort studies. Eur. J. Epidemiol. 26, 863–876.

Johnstone, C.N., Tebbutt, N.C., Abud, H.E., White, S.J., Stenvers, K.L., Hall, N.E., Whitehead, R.H., Cody, S.H., Catimel, B., Nice, E.C., Burgess, A.W., Heath, J.K., 2000. Characterization of mouse A33 antigen, a definitive marker for basolateral surfaces of intestinal epithelial cells. Am. J. Physiol. - Gastrointest. Liver Physiol. 279, 500–510.

Kannarkatt, J., Joseph, J., Kurniali, P.C., Al-Janadi, A., Hrinczenko, B., 2017. Adjuvant Chemotherapy for Stage II Colon Cancer: A Clinical Dilemma. J. Oncol. Pract. 13, 233–241.

Keller, M.S., Ezaki, T., Guo, R., Lynch, J.P., 2004. Cdx1 or Cdx2 expression activates E-cadherin-mediated cell-cell adhesion and compaction in human COLO 205 cells. Am. J. Physiol. Gastrointest. Liver Physiol. 287, G104–G114.

Kim, S., Domon-Dell, C., Wang, Q., Chung, D.H., Di Cristofano, A., Pandolfi, P.P., Freund, J.N., Evers, B.M., 2002. PTEN and TNF-α regulation of the intestinal-specific Cdx-2 homeobox gene through a PI3K, PKB/Akt, and NF-κB-dependent pathway. Gastroenterology 123, 1163–1178.

Kinnersley, B., Labussière, M., Holroyd, A., Di Stefano, A.L., Broderick, P., Vijayakrishnan, J., Mokhtari, K., Delattre, J.Y., Gousias, K., Schramm, J., Schoemaker, M.J., Fleming, S.J., Herms, S., Heilmann, S., Schreiber, S., Wichmann, H.E., Nöthen, M.M., Swerdlow, A., Lathrop, M., Simon, M., Bondy, M., Sanson, M., Houlston, R.S., 2015. Genome-wide association study identifies multiple susceptibility loci for glioma. Nat. Commun. 6, 1–9.

Kinugasa, T., Akagi, Y., Yoshida, T., Ryu, Y., Shiratuchi, I., Ishibashi, N., Shirouzu, K., 2010. Increased claudin-1 protein expression contributes to tumorigenesis in ulcerative colitis-associated colorectal cancer. Anticancer Res. 30, 3181–3186.

Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P.J., Clevers, H., 1998. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. Nat. Genet. 19, 379–383.

Kriegl, L., Horst, D., Reiche, J.A., Engel, J., Kirchner, T., Jung, A., 2010. LEF-1 and TCF4 expression correlate inversely with survival in colorectal cancer. J. Transl. Med. 8, 1–8.

Krishn, S.R., Kaur, S., Smith, L.M., Johansson, S.L., Jain, M., Patel, A., Gautam, S.K., Hollingsworth, M.A., Mandel, U., Clausen, H., Lo, W.C., Fan, W.T.L., Manne, U., Batra, S.K., 2016. Mucins and associated glycan signatures in colon adenoma-carcinoma sequence: Prospective pathological implication(s) for early diagnosis of colon cancer. Cancer Lett. 374, 304–314.

Kuhnert, F., Davis, C.R., Wang, H., Chu, P., Lee, M., Yuan, J., Nusse, R., Kuo, C.J., 2004. Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1. Proc. Natl. Acad. Sci. U. S. A. 101, 266–271.

Landau, M.S., Kuan, S.F., Chiosea, S., Pai, R.K., 2014. BRAF-mutated microsatellite stable colorectal carcinoma: An aggressive adenocarcinoma with reduced CDX2 and increased cytokeratin 7 immunohistochemical expression. Hum. Pathol. 45, 1704–1712.

Larsen, S., Davidsen, J., Dahlgaard, K., Pedersen, O.B., Troelsen, J.T., 2019. HNF4α and CDX2 Regulate Intestinal YAP1 Promoter Activity. Int. J. Mol. Sci. 20.

Läubli, H., Borsig, L., 2019. Altered Cell Adhesion and Glycosylation Promote Cancer Immune Suppression and Metastasis. Front. Immunol. 10, 1–9.

Leedham, S.J., 2014. Measuring stem cell dynamics in the human colon--where there's a wiggle, there's a way. J. Pathol. 234, 292–295.

Li, Chao, Zuo, D., Liu, T., Yin, L., Li, Chenyao, Wang, L., 2019. Prognostic and Clinicopathological Significance of MUC Family Members in Colorectal Cancer: A Systematic Review and Meta-Analysis. Gastroenterol. Res. Pract. 2019.

Liang, K., Zhou, G., Zhang, Q., Li, J., Zhang, C., 2014. Expression of hippo pathway in colorectal cancer. Saudi J. Gastroenterol. 20, 188–194.

Liang, P.S., Chen, T.Y., Giovannucci, E., 2009. Cigarette smoking and colorectal cancer incidence and mortality: Systematic review and meta-analysis. Int. J. Cancer 124, 2406–2415.

Loomans-Kropp, H.A., Umar, A., 2019. Increasing Incidence of Colorectal Cancer in Young Adults. J. Cancer Epidemiol. 2019, 9841295.

Ma, C., Olevian, D., Miller, C., Herbst, C., Jayachandran, P., Kozak, M.M., Chang, D.T., Pai, R.K., 2019. SATB2 and CDX2 are prognostic biomarkers in DNA mismatch repair protein deficient colon cancer. Mod. Pathol. 32, 1217–1231.

Ma, Y., Yang, Y., Wang, F., Zhang, P., Shi, C., Zou, Y., Qin, H., 2013. Obesity and Risk of Colorectal Cancer: A Systematic Review of Prospective Studies. PLoS One 8.

Mallo, G.V., Rechreche, H., Frigerio, J.M., Rocha, D., Zweibaum, A., Lacasa, M., Jordan, B.R., Dusetti, N.J., Dagorn, J.C., Iovanna, J.L., 1997. Molecular cloning, sequencing and expression of the mRNA encoding human Cdx1 and Cdx2 homeobox. Down-regulation of Cdx1 and Cdx2 mRNA expression during colorectal carcinogenesis. Int. J. cancer 74, 35–44.

Mariadason, J.M., Bordonaro, M., Aslam, F., Shi, L., Kuraguchi, M., Velcich, A., Augenlicht, L.H., 2001. Down-regulation of beta-catenin TCF signaling is linked to colonic epithelial cell differentiation. Cancer Res. 61, 3465–3471.

Mattiuzzi, C., Sanchis-Gomar, F., Lippi, G., 2019. Concise update on colorectal cancer epidemiology. Ann. Transl. Med. 7, 609–613.

Melin, B.S., Barnholtz-Sloan, J.S., Wrensch, M.R., Johansen, C., Il'yasova, D., Kinnersley, B., Ostrom, Q.T., Labreche, K., Chen, Y., Armstrong, G., et al., 2017. Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. Nat. Genet. 49, 789–794.

Melo, F.D.S.E., Wang, X., Jansen, M., Fessler, E., Trinh, A., De Rooij, L.P.M.H., De Jong, J.H., De Boer, O.J., Van Leersum, R., Bijlsma, M.F., Rodermond, H., Van Der Heijden, M., Van Noesel, C.J.M., Tuynman, J.B., Dekker, E., Markowetz, F., Medema, J.P., Vermeulen, L., 2013. Poor-prognosis colon cancer is defined by a molecularly distinct subtype and develops from serrated precursor lesions. Nat. Med. 19, 614–618.

Merlos-Suárez, A., Barriga, F.M., Jung, P., Iglesias, M., Céspedes, M.V., Rossell, D., Sevillano, M., Hernando-Momblona, X., Da Silva-Diz, V., Muñoz, P., Clevers, H., Sancho, E., Mangues, R., Batlle, E., 2011. The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. Cell Stem Cell 8, 511–524.

Natoli, M., Christensen, J., El-Gebali, S., Felsani, A., Anderle, P., 2013. The role of CDX2 in Caco-2 cell differentiation. Eur. J. Pharm. Biopharm. 85, 20–25.

Neeman, E., Ben-Eliyahu, S., 2013. Surgery and stress promote cancer metastasis: New outlooks on perioperative mediating mechanisms and immune involvement. Brain. Behav. Immun.

Neumann, J., Heinemann, V., Engel, J., Kirchner, T., Stintzing, S., 2018. The prognostic impact of CDX2 correlates with the underlying mismatch repair status and BRAF mutational status but not with distant metastasis in colorectal cancer. Virchows Arch. 473, 199–207.

Nishiuchi, A., Hisamori, S., Sakaguchi, M., Fukuyama, K., Hoshino, N., Itatani, Y., Honma, S., Maekawa, H., Nishigori, T., Tsunoda, S., Obama, K., Miyoshi, H., Shimono, Y., Mark Taketo, M., Sakai, Y., 2019. Microrna-9-5p-CDX2 axis: A useful prognostic biomarker for patients with stage II/III colorectal cancer. Cancers (Basel). 11.

Olsen, A.K., Coskun, M., Bzorek, M., Kristensen, M.H., Danielsen, E.T., Jørgensen, S., Olsen, J., Engel, U., Holck, S., Troelsen, J.T., 2013. Regulation of APC and AXIN2 expression by intestinal tumor suppressor CDX2 in colon cancer cells.

Carcinogenesis 34, 1361–1369.
Olsen, J., Eiholm, S., Kirkeby, L.T., Espersen, M.L.M., Jess, P., Gögenür, I., Olsen, J., Troelsen, J.T., 2016. CDX2 downregulation is associated with poor differentiation and MMR deficiency in colon cancer. Exp. Mol. Pathol. 100, 59–66.
Olsen, J., Espersen, M.L.M., Jess, P., Kirkeby, L., Troelsen, J.T., 2014. The clinical perspectives of CDX2 expression in colorectal cancer: A qualitative systematic review. Surg. Oncol. 23, 167–176.
Ong, B.A., Vega, K.J., Houchen, C.W., 2014. Intestinal stem cells and the colorectal cancer microenvironment. World J. Gastroenterol. 20, 1898–1909.
Parker, B.C., Zhang, W., 2013. Fusion genes in solid tumors: An emerging target for cancer diagnosis and treatment. Chin. J. Cancer 32, 594–603.
Parr, C., Sanders, A., Jiang, W., 2012. Hepatocyte Growth Factor Activation Inhibitors - Therapeutic Potential in Cancer. Anticancer. Agents Med. Chem. 10, 47–57.
Pilati, C., Taieb, J., Balogoun, R., Marisa, L., Reynie, A. De, 2017. CDX2 prognostic value in stage II / III resected colon cancer is related to CMS classification Original article. Ann. Oncol. 1032–1035.
Piñeros, M., Parkin, D.M., Ward, K., Chokunonga, E., Ervik, M., Farrugia, H., Gospodarowicz, M., O'Sullivan, B., Soerjomataram, I., Swaminathan, R., Znaor, A., Bray, F., Brierley, J., 2019. Essential TNM: a registry tool to reduce gaps in cancer staging information. Lancet Oncol. 20, 103–111.
Pino, M.S., Chung, D.C., 2010. The chromosomal instability pathway in colon cancer. Gastroenterology 138, 2059–2072.
Pinto, D., Gregorieff, A., Begthel, H., Clevers, H., 2003. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. Genes Dev. 17, 1709–1713.
Pinto, R., Hansen, L., Hintze, J., Almeida, R., Larsen, S., Coskun, M., Davidsen, J., Mitchelmore, C., David, L., Troelsen, J.T., Bennett, E.P., 2017. Precise integration of inducible transcriptional elements (PrIITE) enables absolute control of gene expression. Nucleic Acids Res. 45.
Pope, J.L., Ahmad, R., Bhat, A.A., Washington, M.K., Singh, A.B., Dhawan, P., 2014. Claudin-1 overexpression in intestinal epithelial cells enhances susceptibility to adenamatous polyposis coli-mediated colon tumorigenesis. Mol. Cancer 13, 1–13.
Rawat, V.P.S., Humphries, R.K., Buske, C., 2012. Beyond Hox: The role of ParaHox genes in normal and malignant hematopoiesis. Blood 120, 519–527.
Rings, E.H.H.M., Boudreau, F., Taylor, J.K., Moffett, J., Suh, E.R., Traber, P.G., 2001. Phosphorylation of the serine 60 residue within the Cdx2 activation domain mediates its transactivation capacity. Gastroenterology 121, 1437–1450.
Robertson, J.H., Yang, S.Y., Winslet, M.C., Seifalian, A.M., 2009. Functional blocking of specific integrins inhibit colonic cancer migration. Clin. Exp. Metastasis 26, 769–780.
Ryan, É.J., Creavin, B., Khaw, Y.L., Kelly, M.E., Mohan, H.M., Geraghty, R., Ryan, E.J., Kennelly, R., Hanly, A., Martin, S.T., Fennelly, D., McDermott, R., Gibbons, D., O'Connell, P.R., Sheahan, K., Winter, D.C., 2018. Effects of CDX2 on prognosis and chemotherapy responsiveness in mismatch repair-deficient

colorectal cancer. BJS Open 2, 456–463.
Sadanandam, A., Lyssiotis, C.A., Homicsko, K., Collisson, E.A., Gibb, W.J., Wullschleger, S., Ostos, L.C.G., Lannon, W.A., Grotzinger, C., Del Rio, M., Lhermitte, B., Olshen, A.B., Wiedenmann, B., Cantley, L.C., Gray, J.W., Hanahan, D., 2013. A colorectal cancer classification system that associates cellular phenotype and responses to therapy. Nat. Med. 19, 619–625.
Sailaja, B.S., He, X.C., Li, L., 2016. The regulatory niche of intestinal stem cells. J. Physiol. 594, 4827–4836.
Sakaguchi, T., Gu, X., Golden, H.M., Suh, E., Rhoads, D.B., Reinecker, H.C., 2002. Cloning of the human claudin-2 5′-flanking region revealed a TATA-less promoter with conserved binding sites in mouse and human for caudal-related homeodomain proteins and hepatocyte nuclear factor-1α. J. Biol. Chem. 277, 21361–21370.
Sakamoto, N., Feng, Y., Stolfi, C., Kurosu, Y., Green, M., Lin, J., Green, M.E., Sentani, K., Yasui, W., McMahon, M., Hardiman, K.M., Spence, J.R., Horita, N., Greenson, J.K., Kuick, R., Cho, K.R., Fearon, E.R., 2017. BRAFV600E cooperates with CDX2 inactivation to promote serrated colorectal tumorigenesis. Elife 6, 1–25.
Sakuma, K., Aoki, M., Kannagi, R., 2012. Transcription factors c-Myc and CDX2 mediate E-selectin ligand expression in colon cancer cells undergoing EGF/bFGF-induced epithelial-mesenchymal transition. Proc. Natl. Acad. Sci. 109, 7776–7781.
Salari, K., Spulak, M.E., Cuff, J., Forster, A.D., Giacomini, C.P., Huang, S., 2012. CDX2 is an amplified lineage-survival oncogene in colorectal cancer. PNAS Plus 109, 3196–3205.
San Roman, A.K., Aronson, B.E., Krasinski, S.D., Shivdasani, R.A., Verzi, M.P., 2015. Transcription factors GATA4 and HNF4A control distinct aspects of intestinal homeostasis in conjunction with transcription factor CDX2. J. Biol. Chem. 290, 1850–1860.
Satake, S., Semba, S., Matsuda, Y., Usami, Y., Chiba, H., Sawada, N., Kasuga, M., Yokozaki, H., 2008. Cdx2 transcription factor regulates claudin-3 and claudin-4 expression during intestinal differentiation of gastric carcinoma. Pathol. Int. 58, 156–163.
Shigematsu, Y., Inamura, K., Mise, Y., Saiura, A., Rehnberg, E., Yamamoto, N., Ishikawa, Y., Takahashi, S., Kanda, H., 2018. CDX2 expression is concordant between primary colorectal cancer lesions and corresponding liver metastases independent of chemotherapy: A single-center retrospective study in Japan. Oncotarget 9, 17056–17065.
Simmini, S., Bialecka, M., Huch, M., Kester, L., Van De Wetering, M., Sato, T., Beck, F., Van Oudenaarden, A., Clevers, H., Deschamps, J., 2014. Transformation of intestinal stem cells into gastric stem cells on loss of transcription factor Cdx2. Nat. Commun. 5, 1–10.
Slik, K., Turkki, R., Carpén, O., Kurki, S., Korkeila, E., Sundström, J., Pellinen, T., 2019. CDX2 Loss With Microsatellite Stable Phenotype Predicts Poor Clinical Outcome in Stage II Colorectal Cancer. Am J Surg Pathol 43, 1473–1482.
Soda, M., Choi, Y.L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara,

S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y., Mano, H., 2007. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 448, 561–566.

Spit, M., Koo, B.K., Maurice, M.M., 2018. Tales from the crypt: Intestinal niche signals in tissue renewal, plasticity and cancer. Open Biol. 8.

Steinhardt, A.A., Gayyed, M.F., Klein, A.P., Dong, J., Maitra, A., Pan, D., Montgomery, E.A., Anders, R.A., 2008. Expression of Yes-associated Protein, YAP, in Common Solid Tumors. Hum. Pathol. 39, 1582–1589.

Su, W.M., Chen, Z.H., Zhang, X.C., Su, J., Xie, Z., Yan, H.H., Yang, J.J., Chen, H.J., Zhou, Q., Zhong, W.Z., Guo, W.B., Chen, S.L., Wu, Y.L., 2015. Single nucleotide polymorphisms in VTI1A gene contribute to the susceptibility of Chinese population to non-small cell lung cancer. Int. J. Biol. Markers 30, 286–293.

Subtil, C., Guérin, E., Schneider, A., Chenard, M.P., Martin, E., Domon-Dell, C., Duluc, I., Brabletz, T., Kedinger, M., Duclos, B., Gaub, M.P., Freund, J.N., 2007. Frequent rearrangements and amplification of the CDX2 homeobox gene in human sporadic colorectal cancers with chromosomal instability. Cancer Lett. 247, 197–203.

Tabariès, S., Siegel, P.M., 2017. The role of claudins in cancer metastasis. Oncogene 36, 1176–1190.

Tamagawa, Y., Ishimura, N., Uno, G., Yuki, T., Kazumori, H., Ishihara, S., Amano, Y., Kinoshita, Y., 2012. Notch signaling pathway and Cdx2 expression in the development of Barrett's esophagus. Lab. Investig. 92, 896–909.

Tanabe, L.M., List, K., 2017. The role of type II transmembrane serine protease-mediated signaling in cancer. FEBS J. 284, 1421–1436.

Taylor, D.P., Burt, R.W., Williams, M.S., Haug, P.J., Cannon-Albright, L.A., 2010. Population-based family-history-specific risks for colorectal cancer : a Constellation Approach. Gastroenterology 138, 877–885.

Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., Lee, C., Montie, J.E., Shah, R.B., Pienta, K.J., Rubin, M.A., Chinnaiyan, A.M., 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science (80-.). 310, 644–648.

Tong, K., Pellón-Cárdenas, O., Sirihorachai, V.R., Warder, B.N., Kothari, O.A., Perekatt, A.O., Fokas, E.E., Fullem, R.L., Zhou, A., Thackray, J.K., Tran, H., Zhang, L., Xing, J., Verzi, M.P., 2017. Degree of Tissue Differentiation Dictates Susceptibility to BRAF-Driven Colorectal Cancer. Cell Rep. 21, 3833–3845.

Tóth, C., Sükösd, F., Valicsek, E., Herpel, E., Schirmacher, P., Tiszlavicz, L., 2018. Loss of CDX2 gene expression is associated with DNA repair proteins and is a crucial member of the Wnt signaling pathway in liver metastasis of colorectal cancer. Oncol. Lett. 15, 3586–3593.

Tremblay, P.L., Huot, J., Auger, F.A., 2008. Mechanisms by which E-selectin regulates diapedesis of colon cancer cells under flow conditions. Cancer Res. 68, 5167–5176.

Troelsen, J.T., Mitchelmore, C., Spodsberg, N., Jensen, A.M., Norén, O., Sjöström, H.,

1997. Regulation of lactase-phlorizin hydrolase gene expression by the caudal-related homoeodomain protein Cdx-2. Biochem J. 322, 833–838.

Tschaharganeh, D.F., Chen, X., Latzko, P., Malz, M., Gaida, M.M., Felix, K., Ladu, S., Singer, S., Pinna, F., Gretz, N., Sticht, C., Tomasi, M.L., Delogu, S., Evert, M., Fan, B., Ribback, S., Jiang, L., Brozzetti, S., Bergmann, F., Dombrowski, F., Schirmacher, P., Calvisi, D.F., Breuhahn, K., 2013. Yes-associated protein up-regulates jagged-1 and activates the NOTCH pathway in human hepatocellular carcinoma. Gastroenterology 144, 1530-1542.e12.

Tsuge, S., Saberi, B., Cheng, Y., Wang, Z., Kim, A., Luu, H., Abraham, J.M., Ybanez, M.D., Hamilton, J.P., Selaru, F.M., Villacorta-Martin, C., Schlesinger, F., Philosophe, B., Cameron, A.M., Zhu, Q., Anders, R., Gurakar, A., Meltzer, S.J., 2019. Detection of Novel Fusion Transcript VTI1A-CFAP46 in Hepatocellular Carcinoma. Gastrointest. Tumors 6, 11–27.

Van de Wetering, M., Castrop, J., Korinek, V., Clevers, H., 1996. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. Mol. Cell. Biol. 16, 745–752.

van der Bij, G.J., Oosterling, S.J., Beelen, R.H.J., Meijer, S., Coffey, J.C., van Egmond, M., 2009. The perioperative period is an underutilized window of therapeutic opportunity in patients with colorectal cancer. Ann. Surg. 249, 727–34.

van Es, J.H., Haegebarth, A., Kujala, P., Itzkovitz, S., Koo, B.-K., Boj, S.F., Korving, J., van den Born, M., van Oudenaarden, A., Robine, S., Clevers, H., 2012. A Critical Role for the Wnt Effector Tcf4 in Adult Intestinal Homeostatic Self-Renewal. Mol. Cell. Biol. 32, 1918–1927.

Veeman, M.T., Slusarski, D.C., Kaykas, A., Louie, S.H., Moon, R.T., 2003. Zebrafish Prickle, a Modulator of Noncanonical Wnt/Fz Signaling, Regulates Gastulation Movements. Curr. Biol. 13, 680–685.

Verzi, M.., Hatzis, P., Sulahian, R., Philips, J., Schuijers, J., Shin, H., Freed, E., Lynch, J.P., Dang, D.T., Brown, M., Clevers, H., Liu, X.S., Shivdasani, R.A., 2010. TCF4 and CDX2, major transcription factors for intestinal function, converge on the same cis-regulatory regions. Proc. Natl. Acad. Sci. U. S. A. 107, 15157–15162.

Verzi, M.P., Shin, H., San Roman, A.K., Liu, X.S., Shivdasani, R.A., 2013. Intestinal Master Transcription Factor CDX2 Controls Chromatin Access for Partner Transcription Factor Binding. Mol. Cell. Biol. 33, 281–292.

Vogel, L.K., Sæbø, M., Skjelbred, C.F., Abell, K., Pedersen, E.D.K., Vogel, U., Kure, E.H., 2006. The ratio of matriptase/HAI-1 mRNA is higher in colorectal cancer adenomas and carcinomas than corresponding tissue from control individuals. BMC Cancer 6, 1–8.

Wang, H., Burnett, T., Kono, S., Haiman, C.A., Iwasaki, M., Wilkens, L.R., Loo, L.W.M., Van Den Berg, D., Kolonel, L.N., Henderson, B.E., Keku, T.O., Sandler, R.S., Signorello, L.B., Blot, W.J., Newcomb, P.A., Pande, M., Amos, C.I., West, D.W., Bézieau, S., Berndt, S.I., Zanke, B.W., Hsu, L., Lindor, N.M., Haile, R.W., Hopper, J.L., Jenkins, M.A., Gallinger, S., Casey, G., Stenzel, S.L., Schumacher, F.R., Peters, U., Gruber, S.B., Tsugane, S., Stram, D.O., Le Marchand, L., 2014.

Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A. Nat. Commun. 5, 1–7.

Wang, H., Jin, S., Lu, H., Mi, S., Shao, W., Zuo, X., Yin, H., Zeng, S., Shimamoto, F., Qi, G., 2017. Expression of survivin, MUC2 and MUC5 in colorectal cancer and their association with clinicopathological characteristics. Oncol. Lett. 14, 1011–1016.

Williams, B.B., Tebbutt, N.C., Buchert, M., Putoczki, T.L., Doggett, K., Bao, S., Johnstone, C.N., Masson, F., Hollande, F., Burgess, A.W., Scott, A.M., Ernst, M., Heath, J.K., 2015. Glycoprotein A33 deficiency: A new mouse model of impaired intestinal epithelial barrier function and inflammatory disease. DMM Dis. Model. Mech. 8, 805–815.

Xia, J., Zeng, M., Zhu, H., Chen, X., Weng, Z., Li, S., 2018. Emerging role of Hippo signalling pathway in bladder cancer. J. Cell. Mol. Med. 22, 4–15.

Xia, X., Xu, E., Quan, S., Huang, Q., Lai, M., 2009. No association between the polymorphisms in CDX2 coding regions and colorectal cancer in Chinese. Mol. Cell. Biochem. 331, 27–30.

Xu, W., Zhu, Y., Shen, W., Ding, W., Wu, T., Guo, Y., Chen, X., Zhou, M., Chen, Y., Cui, L., Du, P., 2019. Combination of CDX2 expression and T stage improves prognostic prediction of colorectal cancer. J. Int. Med. Res. 47, 1829–1842.

Yan, K.S., Janda, C.Y., Chang, J., Zheng, G.X.Y., Larkin, K.A., Luca, V.C., Chia, L.A., Mah, A.T., Han, A., Terry, J.M., Ootani, A., Roelf, K., Lee, M., Yuan, J., Li, X., Bolen, C.R., Wilhelmy, J., Davies, P.S., Ueno, H., Von Furstenberg, R.J., Belgrader, P., Ziraldo, S.B., Ordonez, H., Henning, S.J., Wong, M.H., Snyder, M.P., Weissman, I.L., Hsueh, A.J., Mikkelsen, T.S., Garcia, K.C., Kuo, C.J., 2017. Non-equivalence of Wnt and R-spondin ligands during Lgr5 + intestinal stem-cell self-renewal. Nature 545, 238–242.

Yoshida, T., Kinugasa, T., Akagi, Y., Kawahara, A., Romeo, K., Shiratsuchi, I., Ryu, Y., Gotanda, Y., Shirouzu, K., 2011. Decreased expression of claudin-1 in rectal cancer: a factor for recurrence and poor prognosis. Anticancer Res. 31, 2517–25.

Yu, J., Liu, D., Sun, X., Yang, K., Yao, J., Cheng, C., Wang, C., Zheng, J., 2019. CDX2 inhibits the proliferation and tumor formation of colon cancer cells by suppressing Wnt/β-catenin signaling via transactivation of GSK-3β and Axin2 expression. Cell Death Dis. 10.

Yu, J., Yu, J., Mani, R., Cao, Q., Brenner, C.J., Cao, X., Wang, G.X., Wu, L., Li, J., Hu, M., Cheng, H., Laxman, B., Vellaichamy, A., Li, Y., Dhanasekaran, S.M., Morey, R., Lonigro, R.J., Tomlins, S.A., Varambally, S., Qin, Z.S., Chinnaiyan, A.M., 2010. An Integrated Network of Androgen Receptor, Polycomb, and TMPRSS2-ERG Gene Fusions in Prostate Cancer Progression. Cancer Cell 17, 443–454.

Zeng, C., Matsuda, K., Jia, W.-H., Chang, J., Kweon, S.-S., Xiang, Y.-B., Shin, A., Jee, S.H., Kim, D.-H., Zhang, B., Cai, Q., et al., 2016. Identification of Susceptibility Loci and Genes for Colorectal Cancer Risk. Gastroenterology 150, 1633–1645.

Zhang, B.Y., Jones, J.C., Briggler, A.M., Hubbard, J.M., Kipp, B.R., Sargent, D.J., Dixon, J.G., Grothey, A., 2017. Lack of Caudal-Type Homeobox Transcription

Factor 2 Expression as a Prognostic Biomarker in Metastatic Colorectal Cancer. Clin. Colorectal Cancer 16, 124–128.

Zhang, J.F., Qu, L.S., Qian, X.F., Xia, B.L., Mao, Z.B., Chen, W.C., 2015. Nuclear transcription factor CDX2 inhibits gastric cancer-cell growth and reverses epithelial-to-mesenchymal transition in vitro and in vivo. Mol. Med. Rep. 12, 5231–5238.

Zheng, J.B., Sun, X.J., Qi, J., Li, S.S., Wang, W., Ren, H.L., Tian, Y., Lu, S.Y., Du, J.K., 2011. Effects of homeodomain protein CDX2 expression on the proliferation and migration of Lovo colon cancer cells. Pathol. Oncol. Res. 17, 743–751.

Zygulska, A.L., Krzemieniecki, K., Pierzchalski, P., 2017. Hippo pathway - brief overview of its relevance in cancer. J. Physiol. Pharmacol. 68, 311–335.